Nature.2023.06.24

Sat, 24 Jun 2023

VISIT…

LANZAROTE Caliente.COM

Nature.2023.06.24

[Sat, 24 Jun 2023]

- This Week
- News in Focus
- Books & Arts
- Opinion
- Work
- Research
- Amendments & Corrections

This Week

- **The world's plan to make humanity sustainable is failing. Science can do more to save it** [20 Jun 2023]
  • There is no planet B, and the UN's Sustainable Development Goals are heading for the rocks. Researchers around the world must do their bit to change that.
- **More carrot, less stick: how to make research assessments fairer** [21 Jun 2023]
  • Research-assessment exercises are often misused to judge researchers or cut their funding — changes to the United Kingdom's scheme are a promising start.

- EDITORIAL
- 20 June 2023

The world's plan to make humanity sustainable is failing. Science can do more to save it

There is no planet B, and the UN's Sustainable Development Goals are heading for the rocks. Researchers around the world must do their bit to change that.



Between now and September, *Nature* will be publishing a special series of editorials covering each of the 17 Sustainable Development Goals.Credit: Stanislav Kogiku/SOPA Images/LightRocket via Getty

Many people would be unable to name even one of the 17 United Nations Sustainable Development Goals (SDGs), which are at the heart of an international project that aims to end poverty and achieve equality while

protecting the environment. From this week, to help raise awareness, we at the Nature Portfolio journals will intensify our ongoing efforts to publish research and commentary on the SDGs.

The SDGs and their 169 associated targets are among humanity's best chance of dealing with global crises, from climate change to economic hardship. World leaders agreed the goals in 2015 and set a 2030 deadline to achieve them. This year, at the half-way point, it looks likely that none of the goals and just 12% of the targets will be met.

In September, world leaders will gather in New York City to come up with a rescue plan. And between now and then, *Nature* will be publishing a series of editorials focusing on the different SDGs, covering what has and hasn't been achieved, what can be done to improve matters, and the part the global scientific community has to play.

The failure to meet even one of the SDGs is not for want of trying. Worldwide, researchers have been aligning their work with the SDGs, along with other global efforts such as UN conventions on climate change and biodiversity loss. Unfortunately, fracturing geopolitics is hindering international cooperation. In addition, there is limited cooperation and coordination across topics and between disciplines.

There is a need to give more consideration to complementarities and trade-offs between the different SDGs. For example, action to develop affordable and clean energy (SDG7) to tackle climate change (SDG13) can have negative local effects on biodiversity (SDGs 14 and 15) through the construction and operation of facilities such as wind and solar farms. And although finance for coal-fired power is an effective way to create work and economic growth (SDG8), it is bad news for health and well-being (SDG3), as well as for the environment. The knowledge of these trade-offs is often given insufficient consideration in policymaking.



[Do the science on sustainability now](#)

Last week, an independent group of science advisers to the UN proposed a way forwards. Their [2023 Global Sustainable Development Report](#) (GSDR) summarizes where the SDGs are failing, and what can be done to rescue them. It reiterates the need for transformational change to get the world onto a sustainable path. Crucially, it recognizes the interconnectedness of the goals and targets.

[Like its predecessor](#), published in 2019, the report recategorizes the SDGs into six “entry points”: human well-being and capabilities; sustainable and just economies; sustainable food systems and healthy nutrition patterns; energy decarbonization with universal access; urban and peri-urban development; and global environmental commons. To progress on human well-being, for example, the report recommends scaling up investment in primary health care and ensuring access to lifesaving interventions; accelerating enrolment in secondary education; and increasing investment in water and sanitation infrastructure.

The authors recognize that the path to sustainability must also include abolishing unsustainable practices, while taking into account the economic and social pain that this can cause. For example, increasing the availability of renewable energy won’t, on its own, tackle climate change: fossil fuels must also be phased out. As we wrote last week (*Nature* **618**, 433; 2023),

there is active resistance to this move and a genuine need to support affected communities, such as those that have relied on the coal industry for decades. Such scenarios don't apply only to reaching energy and climate goals.

The GSDR represents welcome progress on the 'what' of meeting the SDGs. It also proposes what to do on the 'how'. The necessary transformations will be expensive, the authors say — requiring extra annual public and private investment of up to US$2.5 trillion. For efforts to succeed, new ways of governing will be required, with the creation of new institutions and the reform of old ones to put sustainability front and centre. Individual and collective action of the kind already under way will also be needed, but on a bigger scale. And people must be given the right resources and skills to complete the task. This will be especially important in low- and middle-income countries (LMICs).

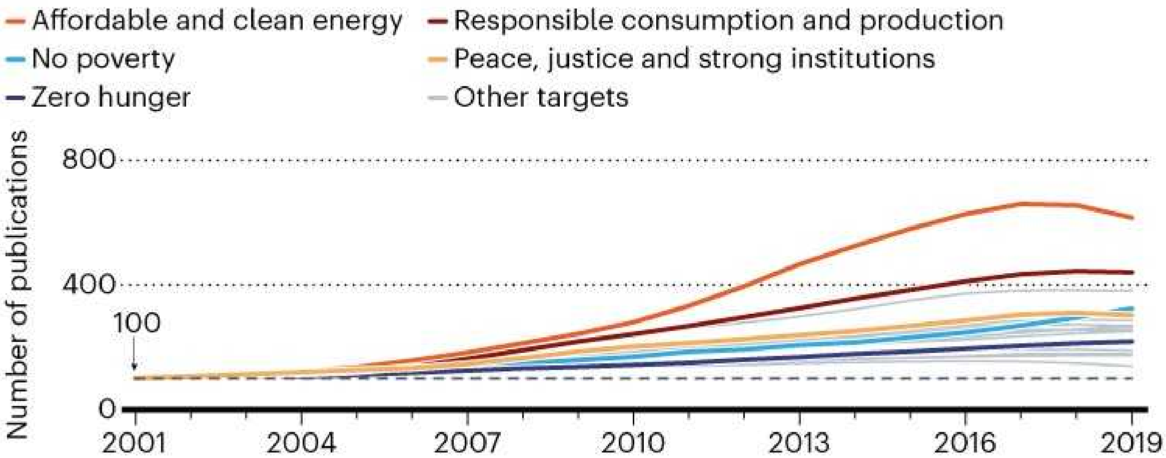
Implicit — and to a degree explicit — in all this is changing how science itself is done. The report argues that the actions that steer the world towards a sustainable path must be rooted in science that is multidisciplinary, equitable and inclusive, openly shared and widely trusted, and "socially robust" — in short, responsive to social context and social needs. As the authors acknowledge, for that to happen, global science needs to evolve. Knowledge needs to be more accessible than it is at present, and the production of that knowledge needs to be more open, too, recognizing, for example, the value of Indigenous and local knowledge to sustainable innovation.

We know from a separate UN study published in 2021 that science in LMICs is already much more aligned with the SDGs than is science in high-income countries. And LMICs have published a much higher volume of research relating to the SDGs (see 'Sustainability science'). The challenge is how to improve the situation in high-income nations. Widespread improvement would be truly game-changing for sustainability.

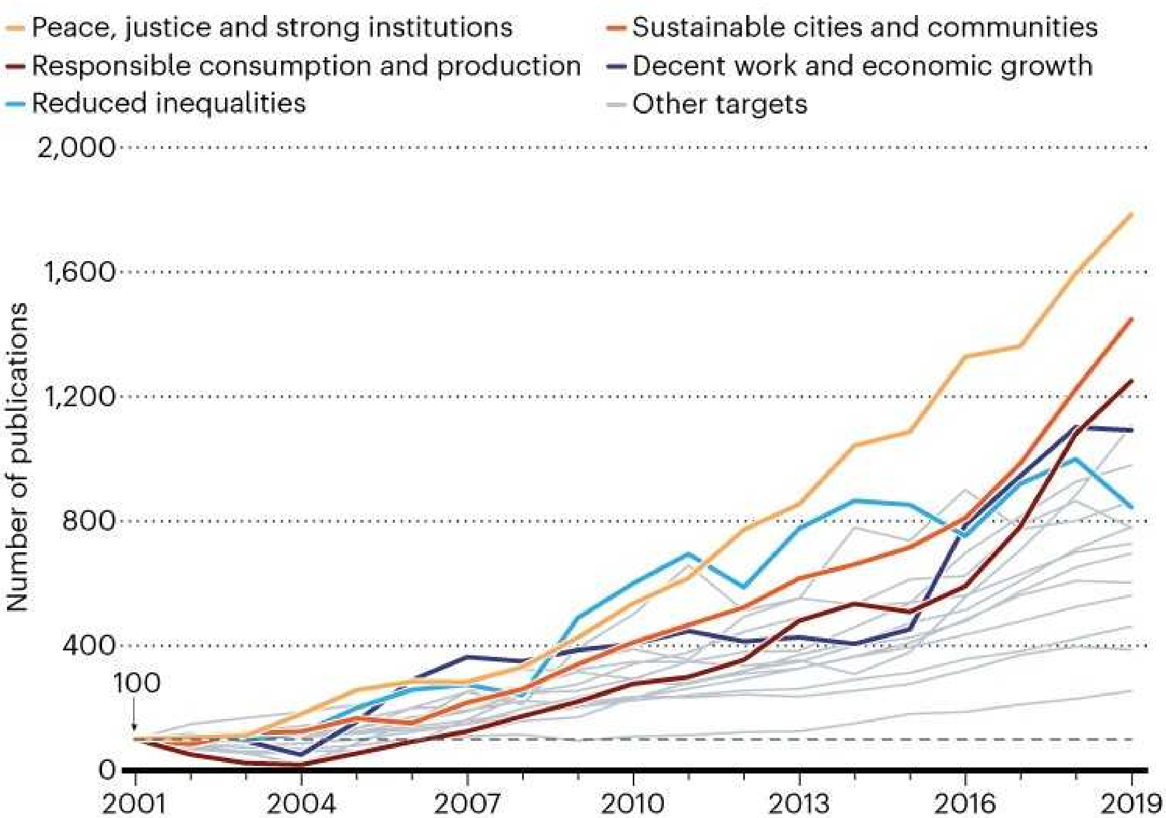
SUSTAINABILITY SCIENCE

Overall, the volume of publications related to the United Nations Sustainable Development Goals increased every year between 2001 and 2019. But there was a sharp contrast in output between high-income countries and the lowest-income countries.

High-income countries



Lowest-income countries



©nature

Adapted from: Changing Directions: Steering Science, Technology and Innovation towards the Sustainable Development Goals

That’s where we’ll take our cue. We’ll assess the evidence and talk to researchers about the state of play on the SDGs, and explore questions that researchers can help to answer. Right now, a sustainable future remains as far away as ever. If there’s even a small chance that we can still achieve the SDGs by 2030, we need to seize it with both hands. As so many have already said, there is no planet B.

Nature **618**, 647 (2023)

doi: https://doi.org/10.1038/d41586-023-01989-9

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01989-9

- EDITORIAL
- 21 June 2023
- Correction 22 June 2023

More carrot, less stick: how to make research assessments fairer

Research-assessment exercises are often misused to judge researchers or cut their funding — changes to the United Kingdom's scheme are a promising start.



National assessment systems tend to be valued more by those in managerial roles, as was the case in some of the first 'time-and-motion' studies of productivity, seen in this image from the 1940s.Credit: Picture Post/Hulton Archive/Getty

How do you measure something such as kindness, teamwork or quality of mentorship? UK researchers now have an opportunity to say what they think. The country's research-funding organizations are planning changes to a nearly 40-year-old system for assessing research so that these qualities are measured and rewarded — and researchers have until 6 October to provide feedback.

The assessments that form part of the UK Research Excellence Framework take place around every seven years. Higher-education institutions are judged on the quality and impact of their research outputs, as well as the quality of the institutional environment.

The results show the proportions of an institution's submitted research projects that reach a standard from 1* (research that is recognized nationally to be original, significant and rigorous) to 4* (world-leading on those criteria). In the latest assessment, in 2021, around 180,000 research outputs from 76,000 researchers at 157 institutions were reviewed by expert panels. More than 80% of submitted research was either 3* or 4* quality.

The REF is a big deal. The results are used to allocate at least £2 billion (US$2.6 billion) in government funding every year, almost all of which rewards institutions for their 4* and 3* research. In a tightening economic climate, the REF provides evidence to both the government and taxpayers that hard-earned money is going to the brightest and the best. For these reasons, the REF is valued more highly by people in managerial positions than it is by researchers, many of whom say that the process creates stress and anxiety, encourages game-playing and discourages novelty.



[Support Europe’s bold vision for responsible research assessment]

To prepare for the REF, university leaders search for what is known as ‘REF-able’ work — research that is most likely to achieve 3* and 4* results. This could mean that 1* or 2* work is less likely to be submitted for assessment even if it is highly relevant to a local issue and done by a research group whose members work well together as a team. That decreases the visibility of these researchers and their work. And less visibility in the REF means less chance of funding.

This is what the research-funding bodies for each of the four UK regions want to change for the next REF exercise in 2028, in line with global trends towards more-responsible research assessment and evaluation. The idea is to take the pressure off individuals, so that the exercise is not used to judge their performance. By extension, universities should not use REF compliance to manage the performance of their staff.

Funding bodies are keen to expand the definition of excellence. They also want all university staff members to be eligible for the assessment, rather than just people directly doing the research, as has been the case so far. In future, contributions to scholarship could also be included from an entry-level technician, or the person at reception who answers the phone. Peer-review panels will also take a more inclusive approach to assessing the quality of outputs. The assessment will take into account not only research

results reported in journals, but also the replication studies, review articles, software, policy reports and the myriad other ways in which academics record their work and its impact. Moreover, the weighting attached to outputs (renamed ‘contribution to knowledge and understanding’) falls from 60% of the overall score to 50%. The environment weighting (renamed ‘people, culture and environment’) increases from 15% to 25%. Impact (renamed ‘engagement and impact’) stays the same at 25%.

External advice

Many of the proposed changes have been informed by the results of several surveys of researchers and advice from an international expert advisory group chaired by Peter Gluckman, New Zealand’s former chief scientific adviser. One of the advisory group’s recommendations, Gluckman says, is for the REF to value research ‘processes’ as much as final results. It means that more journals will need to offer publishing formats, such as Registered Reports, that don’t simply focus on a final outcome.

Peer review will remain the main method for judging quality. But other measures are being considered for evaluating aspects such as data on equality and diversity, with metrics for career progression as proxies for a supportive research culture. There clearly is value in monitoring these data, but it is debatable whether they are a true reflection of a supportive research culture, or just a box-ticking exercise.



Research assessment exercises are necessary — but we need to learn to do them better

Similarly, the ambition that the REF should not be used as a performance-management tool is reassuring. But it's not clear how the changes will take pressure off individuals. Researchers' identities will still need to be known to those doing the assessments. It might be better to explore incentives such as rewarding researchers who submit their work voluntarily. If funding bodies genuinely want to reward a diversity of scholarship, most of the funding should not be going to 4* and 3* research.

Research is as much a craft as it is a science. It takes time to acquire and hone skills. Being a researcher involves continuous learning and improvement. Publishing articles in well-known journals and generating and curating data are not the only signs of scholarship — and being a good mentor to a PhD cohort involves more than passing a set number of graduate students. Research-assessment systems will always be less-than-perfect measures of the totality of the work that happens at universities. The planned reform of the REF is a welcome start. It is an opportunity to make assessment better, deeper and more inclusive, and, ideally, without doing any harm.

Nature **618**, 648 (2023)

doi: https://doi.org/10.1038/d41586-023-01990-2

Updates & Corrections

- **Correction 22 June 2023**: This Editorial gave incorrect weightings for the outputs and impact components of REF 2021. Outputs actually had a weighting of 60%, not 65%, and impact should have been 25%, not 20%.

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01990-2

News in Focus

- **Most advanced synthetic human embryos yet spark controversy** [16 Jun 2023]
  • A pair of studies raises ethical and legal questions about the status of lab-grown human embryo models.
- **How psychedelic drugs achieve their potent health benefits** [14 Jun 2023]
  • Mouse studies suggest that drugs from LSD to ecstasy renew the brain's flexibility — but some scientists are sceptical.
- **Prized dinosaur fossil returned to Brazil after controversy** [14 Jun 2023]
  • The one-of-a-kind specimen will be housed at a museum in Santana do Cariri, near where it was found.
- **Could endometriosis be caused by bacteria? Study offers fresh clues** [14 Jun 2023]
  • Link to bacterial infection suggests a potential way to treat the painful disorder.
- **IBM quantum computer passes calculation milestone** [14 Jun 2023]
  • 'Benchmark' experiment suggests quantum computers could have useful real-world applications within two years.
- **Concussion guidance for sport sidesteps link to brain disease — critics are baffled** [14 Jun 2023]
  • The statement's authors say the connection between repeated head injuries and the condition CTE is based on insufficiently rigorous data.
- **Laos cave fossils prompt rethink of human migration map** [13 Jun 2023]
  • A skull fragment and shin bone suggest that early modern humans might have passed through southeast Asia earlier than thought.
- **Mum's microbes might boost brain development of c-section babies** [15 Jun 2023]
  • Vaginal seeding is safe and seems to benefit infants delivered by the surgery — but larger trials are needed.

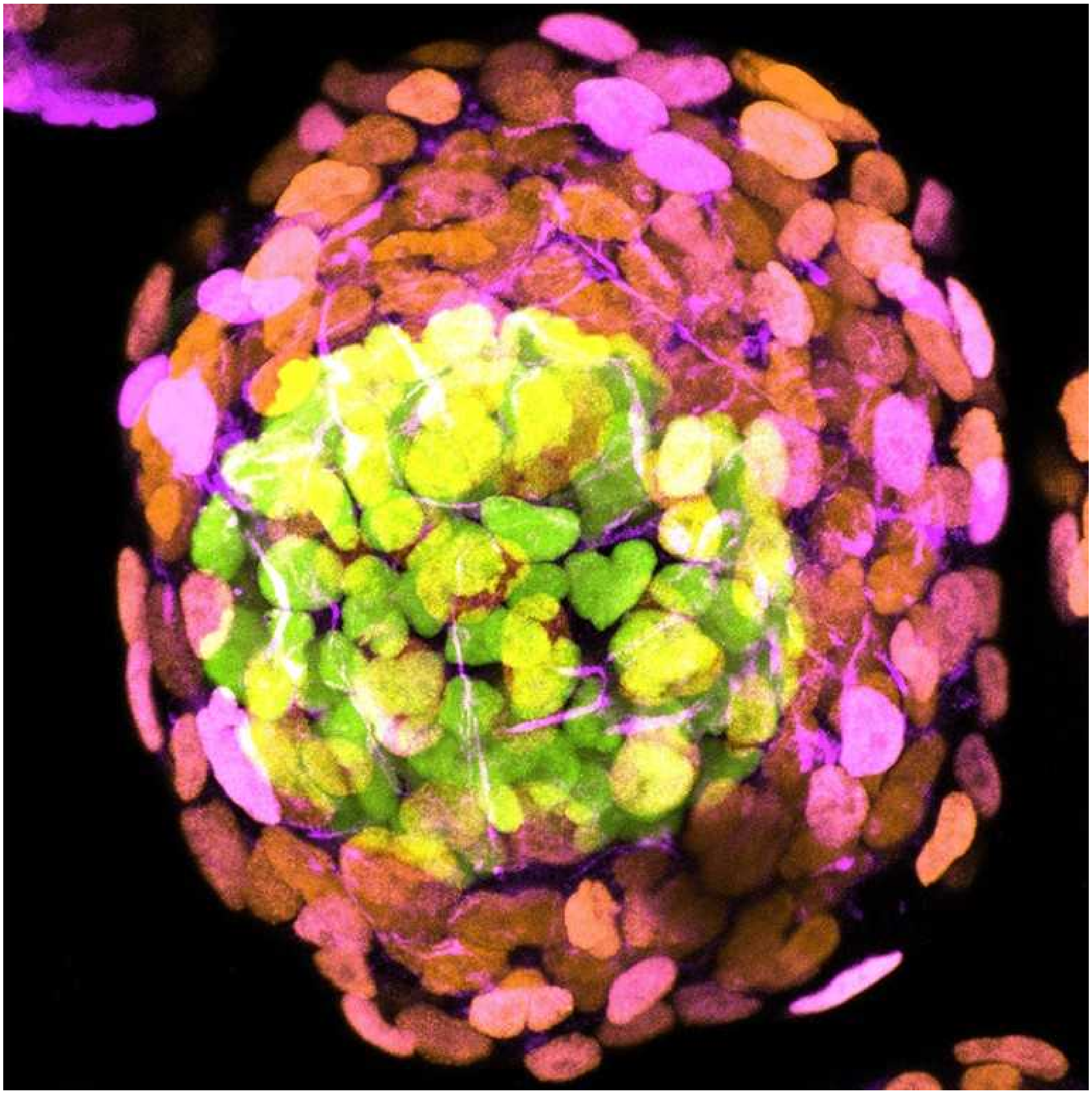
- **Have a high-pitched voice? It might be in your genes** [09 Jun 2023]
  • Survey of nearly 13,000 Icelanders pinpoints for the first time a genetic variant that shapes whether a person’s voice sounds high.
- **Invasive palms and WWII damaged an island paradise. Could fungi help to restore it?** [21 Jun 2023]
  • Researchers ventured to the world’s most remote island to study how fungi in soils could help to revive damaged ecosystems.

- NEWS EXPLAINER
- 16 June 2023
- Correction 20 June 2023

Most advanced synthetic human embryos yet spark controversy

A pair of studies raises ethical and legal questions about the status of lab-grown human embryo models.

- Philip Ball



Embryo-like structures made using human stem cells could enable research that is not currently possible using natural embryos.Credit: Magdalena Zernicka-Goetz, Bailey Weatherbee and Carlos Gantner

Two teams of scientists have announced that they have grown embryo-like structures, made entirely from human stem cells, that are more advanced than any previous efforts. The synthetic embryos developed to a stage equivalent to that of natural embryos about 14 days after fertilization.

Such experiments could provide opportunities to study human embryonic development at later stages than ever before. But they also raise ethical and

legal questions about the status of such ‘embryo models’ and how they should be regulated.

The work is described in two preprint studies[1,2], posted to the bioRxiv server on 15 June by teams led by developmental biologist Magdalena Zernicka-Goetz at the University of Cambridge, UK, and stem-cell biologist Jacob Hanna at the Weizmann Institute of Science in Rehovot, Israel. Both groups had previously presented their findings at scientific meetings, with the work making headlines after Zernicka-Goetz spoke about her results at the annual meeting of the International Society for Stem Cell Research in Boston, Massachusetts, on 14 June.

Nature spoke to scientists about what these developments could mean for research on human embryos.

What did the researchers do and how does it differ from previous work?

Both teams allowed their embryo-like structures to self-assemble from human embryonic stem cells, some of which had been converted into cell types resembling the stem cells that form a placenta and the cells that form the yolk sac outside a naturally developing embryo.

The researchers say that the resulting embryo models show structures and gene transcription profiles found in human embryos between 6 and 14 days after fertilization — up to the onset of the stage called gastrulation, when the cells that will form the embryo become organized into a layer between the amniotic cavity and the yolk sac.

Researchers have made similar entities before from the stem cells of humans and other animals. Last year, both Zernicka-Goetz’s and Hanna’s teams used similar techniques to create embryo models from mouse cells that developed all the way up to the stage at which organs such as the heart and brain begin to form[3,4]. Human embryo models haven’t got that far, but in a preprint posted on bioRxiv on 17 May, stem-cell biologist Ali Brivanlou at the Rockefeller University in New York City and his co-workers reported the development of human embryo models that show signatures of gastrulation

equivalent to those seen at around 12 days after fertilization[5]. The latest studies[1,2] say that they have made the most advanced human embryo models so far.

What is the significance of the embryos lasting for 14 days?

Research on natural human embryos tends to observe a widely adopted guideline — enforced by law in many countries — that human embryos should not be cultured in the laboratory beyond 14 days. This means that researchers have to use animal models to study later stages of embryonic development. These do not necessarily reflect the corresponding processes in humans.

But because in most countries embryo models do not meet the formal definition of an embryo, they are not subject to such restrictions. “We sought to develop a tool to ask specific questions about the second week of human embryo development, since using actual human embryos in research is ethically and technically challenging,” says Zernicka-Goetz.

Models that are older than 14 days could therefore offer important insight into human embryonic development that cannot currently be obtained. They could be used to study developmental defects, for example, or pregnancy loss.

Why is the research scientifically controversial?

Growing embryo models to ever-later stages of development has become a highly competitive race, provoking many arguments about the merits of claims made.

It remains to be seen whether the claims made by the latest studies, neither of which has yet been peer-reviewed, will pass muster. Alfonso Martinez Arias, a developmental biologist at Pompeu Fabra University in Barcelona, Spain, says there is “nothing” in the results described by Zernicka-Goetz and her colleagues that can be considered analogous to real 14-day embryos.

“What we can see is masses of cells separated into compartments, but no embryo-like organization,” he says. He thinks that the over-expression of some genes needed to produce the extra-embryonic cell types “confuses what cells do”, and argues that the results do not show anything that goes beyond earlier work.

Zernicka-Goetz acknowledges the limitations of embryo models for studying development. “These structures do not recapitulate all aspects of the embryo,” she says, “but rather serve as a complementary tool for us to study the differentiation of specific tissues during key stages of development.”

What about ethical concerns?

The results have sparked a discussion about the status of human embryo models in general, and whether they should continue to fall outside legislation on human embryos. Although they are not subject to the 14-day rule, the embryo-like structures reported by Zernicka-Goetz’s and Hanna’s groups do need to respect guidelines and regulations on the use of the human embryonic stem cells from which they are made. But other groups have made embryo models using ‘induced’ stem cells derived from adult tissues[6], which are not governed by such rules. Those embryo models “are not regulated at all”, says Robin Lovell-Badge, a cell biologist at the Francis Crick Institute in London.

So far, no one has made embryo models that have the capacity to develop into human beings, but a recent study on monkey embryo models showed that such models could induce pregnancy (which terminated spontaneously soon after) if placed in the uterus[7].

Some researchers think that a revised definition of an embryo is needed to clarify the issues. For others, the whole purpose of embryo models is to circumvent the current constraints on embryo research. “These models do challenge the need to stick to the 14-day rule,” says Lovell-Badge, who was part of a committee that, in 2021, recommended relaxing the guideline.

In any case, there are significant challenges to making human embryo models that live much longer, says Martinez Arias. Creating structures that

develop up to 21 days “will not be easy”, he says. “I will be surprised if [human embryo models] can go beyond it.”

Nature **618**, 653-654 (2023)

doi: https://doi.org/10.1038/d41586-023-01992-0

Updates & Corrections

- **Correction 20 June 2023**: An earlier version of this article stated that the embryo models show structures and gene transcription profiles found in human embryos during and after gastrulation.

References

1. Weatherbee, B. A. T. *et al.* Preprint at bioRxiv: https://doi.org/10.1101/2023.06.15.545082 (2023).
2. Oldak, B. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2023.06.14.544922 (2023).
3. Tarazi, S. *et al.* *Cell* **185**, 3290–3306 (2022).
4. Amadei, G. *et al.* *Nature* **610**, 143–153 (2022).
5. De Santis, R. *et al.* Preprint at bioRxiv https://doi.org/10.1101/2023.05.16.541017 (2023).
6. Zheng, Y. *et al.* *Nature* **573**, 421–425 (2019).
7. Li, J. *et al.* *Cell Stem Cell* **30**, 362–377 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01992-0

- NEWS
- 14 June 2023

How psychedelic drugs achieve their potent health benefits

Mouse studies suggest that drugs from LSD to ecstasy renew the brain’s flexibility — but some scientists are sceptical.

- Sara Reardon



Crystals of the psychedelic drug MDMA, also called ecstasy, which restores the ability of mice to learn from certain aspects of their environment.Credit: Police Scotland/Contraband Collection/Alamy

Psychedelic drugs are promising treatments for many mental-health conditions, but researchers don't fully understand why they have such powerful therapeutic effects. Now, a study in mice suggests that psychedelics all work in the same way: they reset the brain to a youthful state in which it can easily absorb new information and form crucial connections between neurons[1].

The findings raise the prospect that psychedelic drugs could allow long-term changes in many types of behavioural, learning and sensory system that are disrupted in mental-health conditions. But scientists caution that more research needs to be done to establish how the drugs remodel brain connections.

The study was published on 14 June in *Nature*.

Short trip, long benefits

Psychedelics such as MDMA (also known as ecstasy), ketamine and psilocybin — the active ingredient in magic mushrooms — are known for producing mind-altering effects, including hallucinations in some cases. But each compound affects a different biochemical pathway in the brain during the short-term 'trip', leaving scientists to wonder why so many of these drugs share the ability to relieve depression[2], addiction and other difficult-to-treat conditions in the long term.

Gül Dölen, a neuroscientist at Johns Hopkins University in Baltimore, Maryland, and her colleagues sought answers by studying how psychedelics affect social behaviour in mice. Mice can learn to associate socializing with positive feelings, but only during an adolescent 'critical period', which closes as they become adults.

The scientists trained mice to associate one 'bedroom' in their enclosure with mousy friends and another room with solitude. They could then

examine how psychedelics affected the rodents' room choices — a proxy for whether the drug affects the critical period.

Sociable mice

Dölen's team had previously found[3] that giving MDMA to adult mice in the company of other mice reopened the critical period, making the MDMA-treated animals more likely to sleep in the social room than were untreated mice. This was not surprising: MDMA is well known for promoting bonding in some animals and in humans.

For their new paper, the researchers gave adult mice either MDMA or one of four psychedelic drugs not known to promote sociability: ibogaine, LSD, ketamine and psilocybin. Mice that received any of the psychedelic drugs were more likely to choose the social room than untreated mice, suggesting that each of the drugs could reopen the critical period.



US could soon approve MDMA therapy — opening an era of psychedelic medicine

But mice did not prefer the social room if given enough ketamine to make them unconscious and therefore unaware of their companions. This suggests that the drugs only open the social critical period if they are taken in a social

context. Each drug opened the critical period for a different length of time, ranging from one week for ketamine to more than four weeks for ibogaine.

Next, the team looked at the animals' brains. They found that in certain brain regions, neurons had become more sensitive to the 'love hormone' oxytocin. Dölen suspects that the drugs confer a state called metaplasticity on the neurons, making the cells more responsive to a stimulus such as oxytocin. This state makes them more likely to rewire and form new connections that would indicate the neurons were responding. The neurons also started expressing genes involved in regulating a protein matrix on their surface. Modifying this matrix, Dölen says, could free the neurons' branches to grow and find new connections.

Drugs that hold the key?

Dölen argues that psychedelics function as a master key that can unlock many kinds of critical period — not just one for sociability — by bestowing metaplasticity on neurons. The end result depends on the context in which the drugs were taken: the level of social engagement, in this case. The results indicate, she says, "that there's some mechanistic relationship between critical period opening and that altered state of consciousness that's shared by all psychedelics".



Research round-up: psychedelic medicine

Takao Hensch, a neurologist at Harvard University in Cambridge, Massachusetts, says the paper is “pioneering” in finding biological mechanisms for how psychedelic drugs work. “It gives hope that [critical periods] are not irreversible and a very careful cellular understanding of psychedelic drugs might hold the key to reopening brain plasticity,” he says. He adds that social behaviour is very complex and that the drugs’ effects should be studied in other brain regions.

David Olson, a biochemist at the University of California, Davis, is sceptical. The drugs, he says, could be changing physical connections between neurons in certain parts of the brain, rather than inducing metaplasticity that makes the neurons more open to influence by environmental stimuli.

Dölen is now testing whether the psychedelic drugs can reopen other types of critical period, including those for the motor system. Reopening it, she says, could lengthen the amount of time that people who have had strokes can benefit from physical therapy, which currently works only in the first few months after a stroke.

Nature **618**, 654-655 (2023)

doi: https://doi.org/10.1038/d41586-023-01920-2

Read the related News & Views: ‘Timing is key for behavioural benefits of psychedelics’

References

1. Nardou, R. *et al. Nature* https://doi.org/10.1038/s41586-023-06204-3 (2023).
2. Goodwin, G. M. *et al. N. Engl. J. Med.* **387**, 1637–1648 (2022).
3. Nardou, R. *et al. Nature* **569**, 116–120 (2019).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01920-2

- NEWS
- 14 June 2023

Prized dinosaur fossil returned to Brazil after controversy

The one-of-a-kind specimen will be housed at a museum in Santana do Cariri, near where it was found.

- Meghie Rodrigues



From left: Leonardo Troiano, an archaeologist at the Institute of National Historic and Artistic Heritage in Brasilia; Hermínio de Araújo Júnior, president of the Brazilian Society of Palaeontology, who is based in Rio de

Janeiro; Luciana Santos, minister of science, technology and innovation; Juan Carlos Cisneros, a palaeontologist at the Federal University of Piauí in Teresina; and Aline Ghilardi, a palaeontologist at the Federal University of Rio Grande do Norte in Natal. Credit: Luara Baggi (ASCOM/MCTI)

This celebratory image shows Brazilian researchers and government officials gathered around a prized fossil of the dinosaur *Ubirajara jubatus*. The fossil, newly returned to Brazil from Germany, sparked an outcry in late 2020 after palaeontologists published a now-retracted paper[1] about it. The 110-million-year-old specimen represents the first non-avian dinosaur to be found in South America with structures that seem to be precursors to feathers. None of the authors on the paper was from Brazil, and it was suggested that the removal of the fossil from the country was illegal. A campaign ensued to return the fossil from Germany, where it was being stored in a museum, as part of a movement denouncing colonialism in palaeontology. "I hope that this recognition by Germany will be an inspiration for other countries that hold, in unclear circumstances, other specimens of Brazilian palaeontological biodiversity," said Luciana Santos, Brazil's minister of science, technology and innovation, during a 12 June event in Brasilia at which the fossil was returned, following a long negotiation between countries. It will now be housed at the Plácido Cidade Nuvens Palaeontological Museum in Santana do Cariri, near where the fossil had been found.

Nature **618**, 655 (2023)

doi: https://doi.org/10.1038/d41586-023-01969-z

References

1. Smyth, R. S. H., Martill, D. M., Frey, E., Rivera-Sylva, H. E. & Lenz, N. *Cretac. Res*. https://doi.org/10.1016/j.cretres.2020.104686 (2020).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01969-z

- NEWS
- 14 June 2023

Could endometriosis be caused by bacteria? Study offers fresh clues

Link to bacterial infection suggests a potential way to treat the painful disorder.

- Heidi Ledford



Fusobacterium has been linked with endometriosis.Credit: CNRI/SPL

Infection by a particular group of bacteria could be linked to endometriosis, a painful condition that affects up to 10% of women and girls of reproductive age.

In a study of 155 women in Japan, members of the bacterial genus *Fusobacterium* were found in the uteruses of around 64% of those with endometriosis, and 7% of those who do not have the condition[1]. Follow-up experiments in mice infected with *Fusobacterium* showed that treatment with an antibiotic could reduce the size and frequency of the lesions, which are associated with endometriosis.

Researchers will need to do more work before the findings, published on 14 June in *Science Translational Medicine*, can be used to develop treatments for endometriosis, says Elise Courtois, a genomicist who studies the disease at the Jackson Laboratory in Farmington, Connecticut.

But the results highlight growing interest in the potential role of microbes in endometriosis, which has limited treatment options and whose origins are poorly understood. “There are definitely things that make us suspect the microbiome is implicated in endometriosis,” says Courtois. “Genetics doesn’t explain everything.”

Painful lesions

Endometriosis is a condition in which tissue similar to the lining of the uterus, called the endometrium, grows in other parts of the body — most often organs in the pelvic region. It often causes lesions on reproductive organs and is associated with reduced fertility. Furthermore, people with lesions on their ovaries have an increased risk of developing ovarian cancer.

The most common symptom of endometriosis is pain, which can be severe. Treatments include hormone therapies, which also act as contraceptives, and surgery to remove the lesions. “We want to find new therapies,” says Yutaka Kondo, a cancer biologist at Nagoya University in Japan and a co-author of the paper. “But first we have to know the reason why people suffer from endometriosis.”



It’s time to expand the definition of ‘women’s health’

Kondo and his colleagues analysed endometrial tissue from women with and without the condition. They found that samples from those with endometriosis were more likely to host bacteria belonging to the genus *Fusobacterium*. *Fusobacterium* were often found in the mouth, gut and vagina, and have been linked to other conditions, such as gum disease.

To see whether *Fusobacterium* could affect the course of endometriosis directly, the team transplanted endometrial tissue from one set of mice into the abdominal cavity of another. Within weeks, endometriotic lesions formed in the recipient mice. Using this model, the researchers found that lesions tended to be more abundant and larger in mice that had also been inoculated with *Fusobacterium* than in those that hadn’t. Treating the mice with the antibiotics metronidazole or chloramphenicol, administered vaginally, reduced the development of endometriosis, and shrank the number and size of the lesions.

A clinical trial in women with endometriosis is now under way to find out whether antibiotics could relieve some of their symptoms, says Kondo.

Missing pieces

The results are compelling, but the story is still missing a few key pieces, says Courtois. For example, it would be useful to test the association between *Fusobacterium* and endometriosis in a more diverse population, she says. Courtois and others have successfully lobbied the Connecticut state government to set up a repository of endometrial samples for research that will include tissue from an ethnically diverse population, as well as transgender and gender diverse people. (This article uses ‘women’ to describe people who are at risk for endometriosis, while recognizing that not all people who identify as women have a uterus, and not all people who have a uterus identify as women.)

More extensive research in people is particularly important because mice — which do not menstruate or form spontaneous endometrial lesions — are limited as models of the condition. Kondo’s mouse study focused on lesions that form on the ovaries, whereas in people, lesions can also form elsewhere in the body, including in the colon and bladder. “These are really intriguing results and potentially exciting,” says Krina Zondervan, head of the Nuffield Women’s and Reproductive Health department at the University of Oxford, UK. “But we’re really at the infancy of this.”

Nature **618**, 656 (2023)

doi: https://doi.org/10.1038/d41586-023-01956-4

References

1. Muraoka, A. *et al. Sci. Transl. Med.* **15**, eadd1531 (2023).

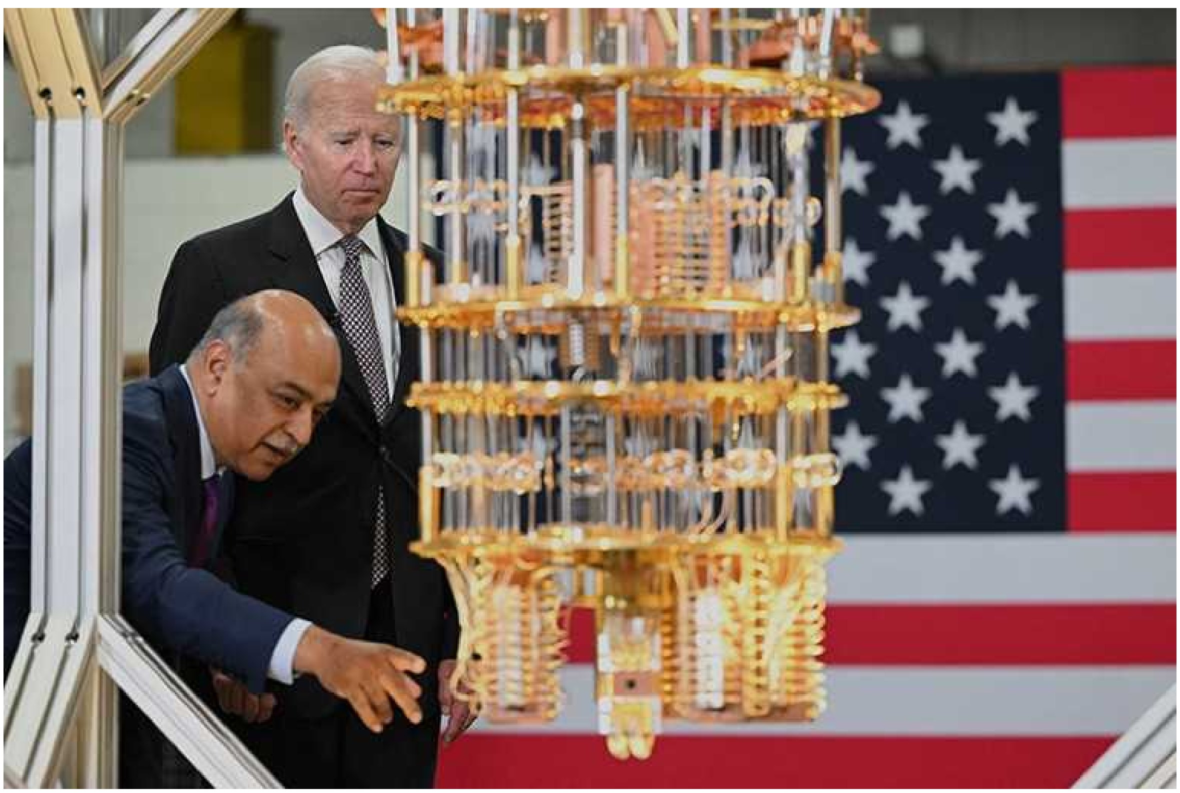
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01956-4

- NEWS
- 14 June 2023

IBM quantum computer passes calculation milestone

‘Benchmark’ experiment suggests quantum computers could have useful real-world applications within two years.

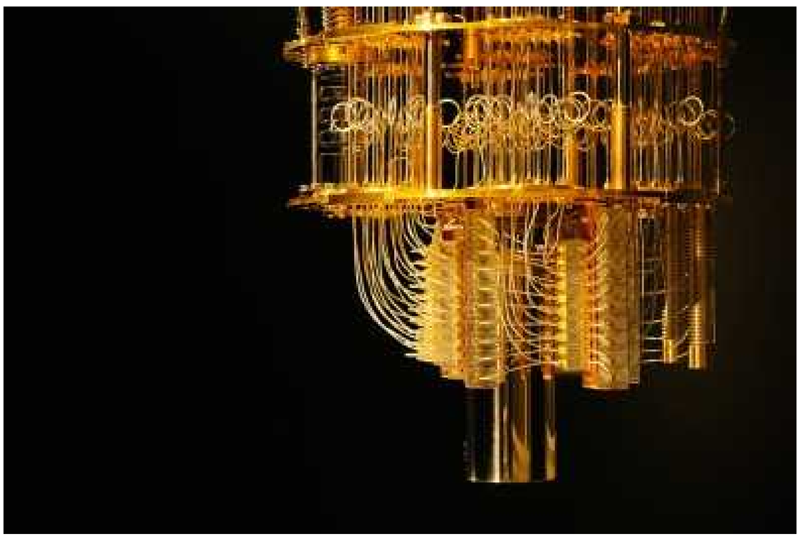
- Davide Castelvecchi



US President Joe Biden and IBM chief executive Arvind Krishna examine a quantum computer at the company’s facility in Poughkeepsie, New York.Credit: Mandel Ngan/AFP via Getty

Four years ago, physicists at Google claimed their quantum computer could outperform classical machines — although only at a niche calculation with no practical applications. Now their counterparts at IBM say they have evidence that quantum computers will soon beat ordinary ones at useful tasks, such as calculating properties of materials or the interactions of elementary particles.

In a proof-of-principle experiment described in *Nature* on 14 June[1], the researchers simulated the behaviour of a magnetic material on IBM's Eagle quantum processor. Crucially, they managed to work around quantum noise — the main obstacle for this technology because it introduces errors in calculations — to get reliable results.



Underdog technologies gain ground in quantum-computing race

Their 'error-mitigating' techniques enabled the team to do quantum calculation "at a scale where classical computers will struggle", says Katie Pizzolato, who heads IBM's quantum theory group in Yorktown Heights, New York.

Although the problem they attacked uses a much-simplified, unrealistic model of a material, "it makes you optimistic that this will work in other systems and more complicated algorithms", says John Martinis, a physicist at the University of California, Santa Barbara, who led the Google team to its 2019 milestone.

Sabrina Maniscalco, chief executive of quantum-computing start-up Algorithmiq in Helsinki, says that the experiment provides a benchmark for the state-of-the-art in quantum computers. “These machines are coming,” she says. Maniscalco’s company is developing algorithms for quantum-chemistry calculations that use error mitigation.

Uniquely quantum

Quantum computers employ peculiarly quantum phenomena, such as the ability of an object to exist in a simultaneous ‘superposition’ of two states, and of multiple objects to share a common quantum state, in what physicists call entanglement. Qubits are the quantum equivalent of the bits of ordinary computers, and can be in a superposition of the ‘0’ and ‘1’ states and be entangled with one another.

Physicists have been experimenting with a range of hardware for building quantum computers, including traps for individual ions or neutral atoms. IBM’s approach — which is also used by Google and other companies — encodes each qubit in a tiny superconducting circuit. For quantum computers to be effective, the qubits have to keep their quantum state for long enough for a calculation to be carried out. So a crucial engineering effort went into increasing the lifetime of the qubits, the IBM team says.



Quantum computers: what are they good for?

In the latest paper, IBM physicist Abhinav Kandala and his collaborators conducted precise measurements of the noise in each of their qubits, which can follow relatively predictable patterns determined by their position inside the device, microscopic imperfections in their fabrication and other factors. Using this knowledge, the researchers extrapolated back to what their measurements — in this case, of the full state of magnetization of a two-dimensional solid — would look like in the absence of noise. They were then able to run calculations involving all of Eagle's 127 qubits and up to 60 processing steps — more than any other reported quantum-computing experiment.

Error approach

Martinis says that the results validate IBM's short-term strategy, which aims to provide useful computing by mitigating, as opposed to correcting, errors. Over the longer term, IBM and most other companies hope to shift towards quantum error correction, a technique that will require large numbers of additional qubits for each data qubit. (Google's strategy has focused on refining quantum error-correction techniques.)

Some researchers are less optimistic about the potential of noise mitigation, and expect that only quantum error correction will enable calculations that would be impossible on even the largest classical supercomputers[2].

The Eagle has 127 qubits — but IBM expects to unveil its most powerful processor yet, the 1,121-qubit Condor chip, later this year. The company also has "utility-scale processors" with up to 4,158 qubits in its development pipeline, says Jay Gambetta, head of IBM's quantum-technology efforts. He adds that to achieve the longer-term goal of building 100,000-qubit machines that can do fully error-corrected algorithms by 2033, researchers will need to solve substantial engineering problems.

Nature **618**, 656-657 (2023)

doi: https://doi.org/10.1038/d41586-023-01965-3

References

1. Kim, Y. *et al. Nature* **618**, 500–505 (2023).

2. Daley, A. J. *et al. Nature* **607**, 667–676 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01965-3

- NEWS
- 14 June 2023

Concussion guidance for sport sidesteps link to brain disease — critics are baffled

The statement’s authors say the connection between repeated head injuries and the condition CTE is based on insufficiently rigorous data.

- Katharine Sanderson



New Zealand’s Portia Woodman tackles South Africa’s Chane Stadler in a rugby sevens match.Credit: William West/AFP via Getty

An influential team of researchers has updated the scientific consensus on how concussion in sport should be defined, treated and monitored. But critics say that the statement, which is revised every four to five years, excludes evidence that links head injuries in sport with long-term brain conditions such as chronic traumatic encephalopathy, or CTE — a high-profile issue in games such as American football and soccer.

The consensus statement, compiled by 114 co-authors after the International Conference on Concussion in Sport in Amsterdam last October, summarizes the latest evidence on sport-related concussion to help clinicians manage the trauma. The latest version introduces details including a description of brain-chemistry events that happen after a concussion. It was published in the *British Journal of Sports Medicine* on 14 June[1].



Why sports concussions are worse for women

But some researchers have criticized the authors' work. "Their refusal to acknowledge a causal relationship between contact-sports participation and CTE is a danger to the public," says Chris Nowinski, a neuroscientist and chief executive of the Concussion Legacy Foundation in Middletown, Delaware, which supports athletes and veterans affected by concussion and CTE.

Many studies have linked repeated sport-related head injuries with CTE — a degenerative brain disease that can develop into dementia. But the consensus authors say that these studies use data from brain banks — where former athletes donate their tissue to be studied after death — that are not rigorous enough to be included in the review.

Data disagreement

“The CTE literature is almost exclusively case-series studies,” says Bob Cantu, a co-author of the consensus report and a neurosurgeon at the Boston University School of Medicine in Massachusetts. “And that literature did not meet the inclusion criteria for the systematic review.”

“I think it is a legitimate criticism of many of the case-series studies that they didn’t look at other confounding factors,” says Cantu. These include a broad range of considerations, from a person’s educational attainment, socio-economic status and diet, to whether they have cardiovascular disease or diabetes.

The authors’ review concluded that amateur athletes are not at greater risk of developing neurodegenerative diseases than is the general population. However, they say that studies of professional American football and soccer players show greater mortality rates from the neurodegenerative disease amyotrophic lateral sclerosis, or ALS. (A link between ALS and CTE has been suggested but is not clear.)

Nowinski says that the link between CTE and concussion is evident. “CTE is caused by repeated head impacts and most frequently seen in the context of contact-sports participation,” he says. “The majority of individuals being exposed to repetitive head impacts are children. We encourage sports organizations to listen to the independent experts and begin to aggressively take steps to prevent CTE. Failure to act will cause more athletes to develop CTE, and more families to suffer,” Nowinski adds.

Marshall Garrett, a legal-medicine specialist at the University of Glasgow, UK, says that the report recommends funding long-term studies into links

between concussion and CTE, but ignores evidence that has been available for years.

Nevertheless, Garrett is delighted to see the addition of detail in the definition of concussion that describes a brain-chemistry event called the ‘neurotransmitter cascade’, which follows a traumatic brain injury. The force to the head triggers a series of events that stop biochemical signals through the brain from working properly. “There are some very good things,” he says of the consensus.

After injury

The statement also updates the protocols for athletes to return to playing sport following a concussion, which Garrett also welcomes. “I’m delighted to see that there is a good definition of return to play,” he says. But he cautions that the time frame for return covers both professional and amateur players and does not necessarily tally with the recommendation in local sports that people should wait 21 days before returning to play after a concussion.

Two gaps in the concussion data include those on children aged 5–12 years and on how the experience varies for men and women.

The statement also doesn’t do enough to address sex differences, says Liz Williams, who studies biomechanics and head injuries at Swansea University, UK. Studies show that female athletes experience concussion more frequently than do men: a 2021 study[2] of around 80,000 US school soccer players showed that girls were concussed almost twice as often as boys.

“This totally fails women by failing to look at them as a potentially separate entity with a different neurophysiological recovery pattern,” says Garrett.

“A lot of the literature is not focused on female athletes and so this is an important consideration for the future,” says Kathryn Schneider at the Sport Injury Prevention Research Centre, University of Calgary, Canada, and joint first author on the consensus statement. “In our systematic review that

evaluated prevention strategies, only 6 out of 220 studies were focused on female athletes,” she says. “We need more research.”

Nature **618**, 657-658 (2023)

doi: https://doi.org/10.1038/d41586-023-01964-4

References

1. Patricios, J. S. *et al. Br. J. Sports Med*. https://doi.org/10.1136/bjsports-2023-106898 (2023).
2. Bretzin, A. C., Covassin, T. & Wiebe, D. J. *JAMA Netw. Open* **4**, e218191 (2021).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01964-4

- NEWS
- 13 June 2023
- Correction 14 June 2023

Laos cave fossils prompt rethink of human migration map

A skull fragment and shin bone suggest that early modern humans might have passed through southeast Asia earlier than thought.

- Jude Coleman



Researchers laboriously sifted through clay, bucket by bucket, using their fingers to hunt for bone fragments. Credit: Fabrice Demeter

Archaeologists have uncovered two new bone fragments in a cave in northern Laos, suggesting that *Homo sapiens* wandered southeast Asia up to 86,000 years ago. The findings, published this week in *Nature*

Communications[1], indicate that humans migrated through the area earlier than previously thought.

Over more than a decade, excavations in the Tam Pà Ling cave have uncovered seven bone fragments sandwiched between layers of clay. Laura Shackelford, an anthropologist at the University of Illinois Urbana–Champaign, and her colleagues have regularly had to hike through sticky tropical heat to reach the mountain-top cave.

After digging 7 metres down, excavations have finally hit bedrock, and the team has been able to reconstruct a complete chronology of the cave, says Shackelford. Sediment and bones unearthed in the cave show that modern humans have inhabited the mountainous region for at least 68,000 years, and passed through even earlier.

“I can’t overestimate the importance of having another point on our map for early modern humans in southeast Asia,” says Miriam Stark, an anthropological archaeologist at the University of Hawaii at Manoa who was not involved in the work. “Understanding southeast Asia is critical to understanding the world’s deep history,” she says.

Old bones

The newly discovered bones are a small piece of skull and a shin bone fragment. The remains were probably swept into the uninhabited cave during a flood. The researchers dated herbivore teeth found alongside the human fossils using electron spin resonance and uranium-series dating. They also estimated the age of the cave sediment using luminescence dating, which calculates when photons last illuminated the soil. Together, the results put the ages of the skull fragment and shin bone at around 70,000 and 77,000 years old, respectively. But the shin bone could be as old as 86,000 years. That’s much older than the first fossil unearthed at the site more than a decade ago, a chunk of a skull estimated to be 46,000 years old[2]. It’s also older than other cave bones — two jaw fragments, a rib and a foot bone — that are between 46,000 and 70,000 years old[3–5].

Fossil records in southeast Asia are limited, in part because the tropical climate decomposes most bones. The details of when early humans first arrived in the region, where they came from and where they migrated to is still a matter of debate, Shackelford says. Laos lies on a potential migration route to Australia, where the oldest archaeological site is around 65,000 years old[6]. In addition to contributing data to an understudied area, Tam Pà Ling provides further insights into the timing of migration through the region.

Some human migration hypotheses use DNA analysis to argue that *H. sapiens* dispersed in a single rapid event after a geological period, called Marine Isotope Stage 5, which lasted from 130,000 to 80,000 years ago. But the Tam Pà Ling fossils don't align with these models. Instead, the fossils suggest dispersal occurred before Marine Isotope Stage 5 had concluded.

"We're seeing something different," Stark says. That doesn't mean that the genetic models are incorrect, she says, just that the picture they reveal is incomplete.

The shape of the Tam Pà Ling fossils further complicates the story. Although they are from *H. sapiens*, the youngest bone — the 46,000-year-old skull fragment — has a mixture of characteristics of both archaic and modern humans, whereas the oldest fossils have more-modern features. For example, the older skull fragment lacks the pronounced brow bone associated with more-archaic humans that is observed to some degree in the younger fossil.

That is counter-intuitive, Shackelford says, and suggests the older fossils might not have evolved from local populations, but rather represent groups of early modern humans that migrated through the area. Armand Mijares, an archaeologist at the University of the Philippines Diliman in Quezon City, says that this is a plausible interpretation, but that more evidence is needed to be certain.

Shackelford and her colleagues will continue excavations in the cave to look for more fossils. They are also trying to recover environmental DNA from the clay, which could provide clues about what flora and fauna were living in the surrounding area tens of thousands of years ago. Discoveries beyond

the cave could also yield valuable insights into the early human inhabitants of the region.

“This is just one mountain,” says Shackelford. “There are thousands of caves to explore.”

Nature **618**, 658-659 (2023)

doi: https://doi.org/10.1038/d41586-023-01903-3

Updates & Corrections

- **Correction 14 June 2023**: An earlier version of this story misstated how far down researchers excavated in the cave.

References

1. Freidline, S. E. *et al. Nature Commun.* https://www.nature.com/articles/s41467-023-38715-y (2023).
2. Demeter, F. *et al. Proc. Natl Acad. Sci. USA* **109**, 14375–14380 (2012).
3. Shackelford, L. *et al. Quat. Int.* **466**, 93–106 (2018).
4. Demeter, F. *et al. PLoS ONE* **10**, e0121193 (2015).
5. Demeter, F. *et al. Curr. Anthropol.* **58**, S527–S538 (2017).
6. Clarkson, C. *et al. Nature* **547**, 306–310 (2017).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01903-3

- NEWS
- 15 June 2023

Mum’s microbes might boost brain development of c-section babies

Vaginal seeding is safe and seems to benefit infants delivered by the surgery — but larger trials are needed.

- Ewen Callaway



Babies born by c-section can be smeared with their mother’s vaginal fluids.Credit: APHP-Cochin-Voisin/Phanie/SPL

Swabbing caesarean-section babies with their mum’s vaginal fluids to encourage the development of a microbiome seems to be safe — and might improve brain development, finds the biggest-yet study of the controversial practice[1].

The randomized and blinded trial of 76 infants and mothers found that babies who had been swabbed with vaginal fluids from their mothers achieved a slightly higher number of neurodevelopmental milestones by three and six months than did c-section babies who did not receive the treatment. But it is not clear whether these differences will be meaningful or lasting, say scientists.



Could baby's first bacteria take root before birth?

"This is not going to make a difference between going to Harvard or not," says Jose Clemente, a microbiologist at Icahn School of Medicine at Mount Sinai in New York and a co-author of the study, which was conducted at a hospital in China. "It's up to future studies to determine exactly what the mechanism is and how can we maximize the benefit."

Microbiome's birth

Mothers who deliver their babies naturally provide them with the beginnings of a microbiome — the body's collection of microbes in the gut, on skin and elsewhere. Vaginal seeding involves collecting a swab of vaginal fluids from the mother and smearing them soon after birth on the skin of a baby born by c-section, with the aim of replicating this first exposure.

Studies have identified differences in the microbiomes of c-section babies and those delivered vaginally[2]. C-section babies had higher levels of opportunistic bacteria — those that circulate in hospital — in their guts in the days after birth, and, months later, tended to lack common gut microbes that would support immune function.

In 2016, a team co-led by Clemente published a trial of vaginal seeding: four infants who underwent the procedure successfully acquired microbes from their mother's swabs and had microbiomes that resembled those of children delivered vaginally[3].

Although the study did not look at the long-term effects of vaginal seeding, the findings led to an avalanche of interest in the procedure, says Clemente. In 2017, the American College of Obstetricians and Gynecologists expressed concern that vaginal seeding could transmit pathogens and recommended that the procedure be conducted only as part of clinical trials[4].

Developmental advances

To better understand the safety of vaginal seeding and determine its potential benefits, a team led by clinical scientist Yan He at Southern Medical University in Guangzhou, China, recruited women who were due to deliver by c-section and randomly selected them to have their newborns exposed to either vaginal swabs or sterile saline.

Women were first screened for numerous infectious diseases that could threaten the health of the babies — including COVID-19, because the trial was conducted at the height of the pandemic. Babies who received the vaginal swabs developed no serious health problems, and had similar rates of non-serious complications, including mild skin conditions and fever, as the babies swabbed with saline.

When the babies were three then six months old, the researchers asked their parents to fill out a checklist of neurodevelopmental milestones in communication, movement, problem-solving and social and personal skills, such as reaching out for toys and smiling at their reflections in a mirror. The

infants who received the vaginal swabs tended to reach slightly more of their milestones than did babies swabbed with saline.

“It’s very exciting and promising,” says Alexander Khoruts, a physician-scientist at the University of Minnesota in Minneapolis. But it’s not clear whether the neurodevelopmental differences will turn out to be meaningful, because the brain changes so much as children develop. “Early delay by a couple months may not translate into anything meaningful at 18 years.”

Lars Engstrand, a clinical microbiologist at the Karolinska Institute in Stockholm, is reassured by the safety of vaginal seeding, but only if the mothers are screened for pathogens. “That is a really important message to clinical researchers in this field.”

Encouraging results

He agrees that his team’s results should be taken with a grain of salt. “I don’t think our data now has real clinical significance to say ‘yes, you should do this intervention’,” he says. Rather, he says it gives him the confidence to embark on a larger and longer-term study that will ask whether vaginal seeding affects rates of neurodevelopmental conditions such as attention deficit hyperactivity disorder and autism spectrum disorder, which population-health studies suggest are more common in people born through c-section.

Other clinical trials are also trying to pinpoint potential benefits of vaginal seeding. Engstrand and his colleagues are measuring its effect on a skin condition called atopic dermatitis in around 300 participants. Clemente is also part of a US study that is looking at allergies and asthma, and a team led by Maria Gloria Dominguez-Bello, a microbial ecologist at Rutgers University in New Brunswick, New Jersery, is looking at its effects on obesity.

Vaginal swabs are not the only way to alter the microbiomes of c-section babies: a 2020 study found that giving newborns diluted faecal samples from their mothers led to gut microbiomes that resembled those of children delivered vaginally[5]. Khoruts says that such observations should stimulate

researchers to identify the specific microbes transferred during birth. “That’s what the next steps have to be.”

Nature **618**, 659-660 (2023)

doi: https://doi.org/10.1038/d41586-023-01972-4

References

1. Zhou, L. *et al. Cell Host Microbe* https://doi.org/10.1016/j.chom.2023.05.022 (2023).
2. Shao, Y. *et al. Nature* **574**, 117–121 (2019).
3. Dominguez-Bello, M. G. *et al. Nature Med.* **22**, 250–253 (2016).
4. Committee on Obstetric Practice. *Obstet. Gynecol.* **130**, e274–e278 (2017).
5. Korpela, K. *et al. Cell* **183**, 324–334 (2020).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01972-4

- NEWS
- 09 June 2023
- Correction 14 June 2023

Have a high-pitched voice? It might be in your genes

Survey of nearly 13,000 Icelanders pinpoints for the first time a genetic variant that shapes whether a person’s voice sounds high.

- Freda Kreier



Icelandic singer-songwriter Björk is known for her unusual and ethereal voice.Credit: Santiago Felipe/ABA/Getty

Scientists have long suspected that a person’s genetics are involved in shaping how their voice sounds. But knowing which genes are involved — and how they wield their influence — has remained out of reach. Now, what researchers are calling the first large-scale genetic study of human voices has identified mutations in a gene that correlate with higher voice pitch, regardless of a person’s sex. It was published on 9 June in *Science Advances*[1].

Before this study, researchers had pinpointed genetic effects on the voice mainly by identifying mutations that lead to speech disorders in people. But “this is the first time we know of a genetic variant that is really, truly affecting voice pitch” across a large population, says Rósa Signý Gísladóttir, a geneticist and linguist at deCODE Genetics, the genomics company in Reykjavik that carried out the investigation.

A thorough check-up

The human voice is partially shaped by body size and hormones: these factors are often responsible for why female voices tend to have a higher pitch than do male voices. But another factor is genetics.

“It’s common knowledge that voices tend to travel in families,” says Kári Stefánsson, a geneticist and the chief executive of deCODE. Aside from anecdotally observing this hereditary link, researchers have conducted a handful of studies to confirm it, including investigations of twins[2,3], in which genetic influences could be separated from environmental ones.



[Baby bats babble like human infants]

But determining which genes play a part requires sampling a lot of people — and therefore a lot more leg work. That's where deCODE comes in. The company, founded in 1996, has mined a vast trove of genetic information from the people of Iceland, making the country a hotspot for genetics research. Because many of the people who live there can trace their family trees back to a small number of ancestors, and because the country has kept detailed genealogical records, it is easier than usual to spot variants and correlate them with traits or diseases.

More than half of Iceland's adult population has now participated in one of deCODE's genetic studies, according to the firm's website. In exchange for free health information, volunteers undergo four hours of intensive observation on everything from their bone density to their mental health, Stefánsson says. "You could basically look at it as an extraordinarily thorough check-up," he adds.

For the voice-pitch study, deCODE recorded the speech of nearly 13,000 Icelanders, and then compared the frequencies of their voices with its genetic database. The team found that shared mutations in one gene — called *ABCC9* — correlated with speaking at a higher pitch, across the age and sex spectrum.

Biological mechanism

Exactly how genetics is causing this phenomenon isn't clear, but the team has some hypotheses. Among other things, *ABCC9* contains instructions for making an ion channel that helps to ensure the proper functioning of the proteins collagen and elastin. These proteins help tissues in the body to stretch — a must for vocal cords to vibrate.



[Why is COVID life-threatening for some people? Genetics study offers clues]

Furthermore, the researchers found that people with high-voice-pitch mutations were more likely to have high pulse pressure — a risk factor for cardiovascular disease. On the surface, this might seem unusual. But it makes sense, the team says, because collagen and elastin also have a role in the stiffness of blood vessels. If mutations to *ABCC9* can alter blood-vessel properties, it's not a stretch to think they could alter vocal properties, too, Gísladóttir says.

Stefánsson says that he thinks this isn't the end of the story, and that other genes are probably involved in shaping voice pitch. To find them, however, even more statistical power — volunteers — is needed.

Large-scale genetic studies peering into human voices like this one are "sorely needed", says Julie Miller, a neuroscientist at the University of Arizona in Tucson. She adds that deCODE's study shows a nice correlation between *ABCC9* mutations and voice pitch, but to prove that the mutations actually cause a higher pitch will require testing in animals.

Gísladóttir is excited by the prospect of exploring this further. The discovery "is truly giving us something new about the biological mechanism of voice that we did not know before".

Nature **618**, 660 (2023)

doi: https://doi.org/10.1038/d41586-023-01901-5

Updates & Corrections

- **Correction 14 June 2023**: An earlier version of this story said that people with high-voice-pitch mutations were more likely to have heart conditions. They are more likely to have high pulse pressure, which is a risk factor for cardiovascular disease. The researchers think the mutations could not only be altering the stiffness of blood vessels but also vocal cords.

References

1. Gisladottir, R. S. *et al. Sci. Adv.* **9**, eabq2969 (2023).

2. Debruyne, F., Decoster, W., Van Gijsel, A. & Vercammen, J. *J. Voice* **16**, 466–471 (2002).

3. Van Lierde, K. M., Vinck, B., De Ley, S., Clement, G. & Van Cauwenberge, P. *J. Voice* **19**, 511–518 (2005).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01901-5

Lost world

Invasive palms and WWII damaged an island paradise.

Could fungi help to restore it?

By Virginia Gewin

21 June 2023

Res
dan

Thi

Leg
squ
carı

Evo

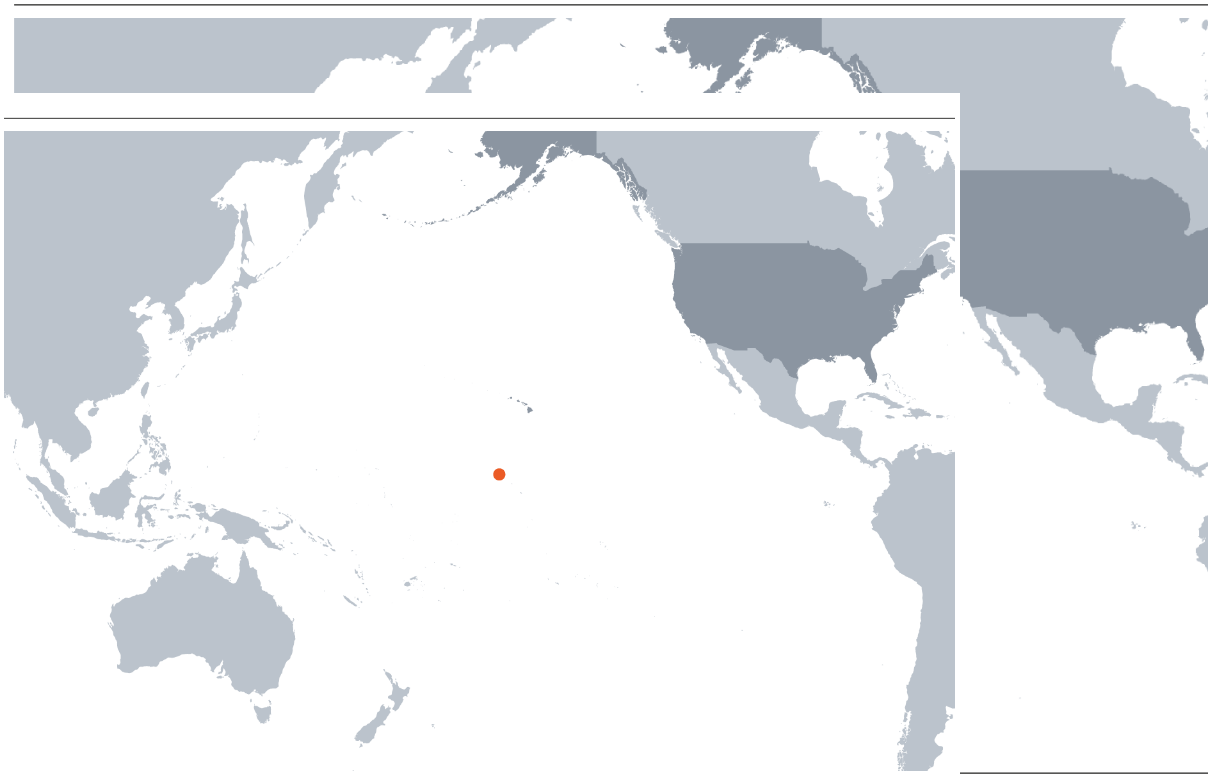
Cre

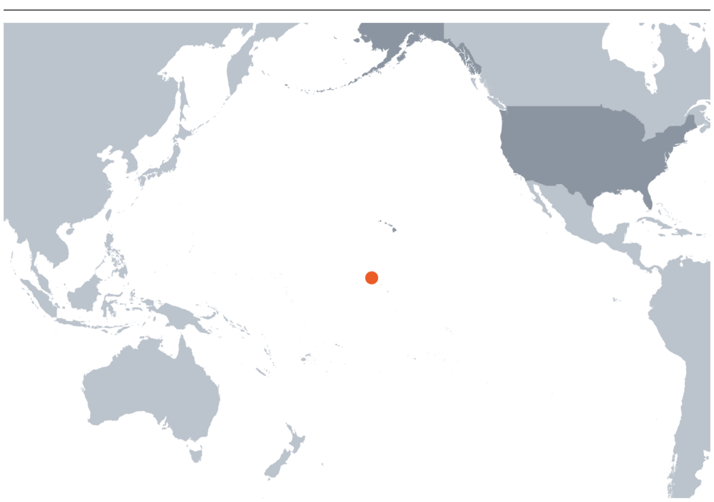
Kie
glol
lool
help
dov



Credit: Toby Kiers/SPUN

Credit: Toby Kiers/SPUN





Nature publications remain neutral with regard to contested jurisdictional claims in published maps.

States
ific

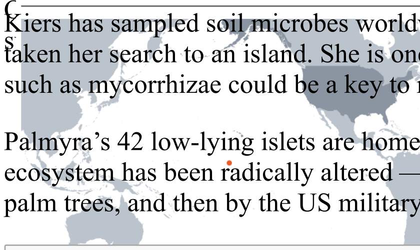
an

vaii

a atoll

Remote WONDER

Located in the centre of the Pacific



Kiers has sampled soil microbes worldwide, but her trip to Palmyra in November 2022 was the first time she had taken her search to an island. She is one of several researchers exploring whether below-ground microorganisms such as mycorrhizae could be a key to restoring heavily degraded ecosystems on these and other islands.

Palmyra's 42 low-lying islets are home to some of the most pristine coral reefs in the world, but the above-water ecosystem has been radically altered — first by nineteenth-century coconut plantations that introduced invasive palm trees, and then by the US military, which dredged areas to create airstrips during the Second World War.

with regard to contested jurisdictional claims in published maps.

Rich coral reefs surround Palymra atoll. (This video has no sound.) Credit: The Nature Conservancy

ited

Rich coral reefs surround Palymra atoll. (This video has no sound.) Credit: The Nature Conservancy

tes

cific

ean

waii

Now the atoll is uninhabited, except for a rotating cast of researchers testing how to repair degraded ecosystems. The non-profit global environmental organization The Nature Conservancy (TNC) bought Palmyra for US$30 million in 2000 and later sold part to the US government. The atoll is co-managed by TNC and the US Fish and Wildlife Service. The two organizations, along with Island Conservation, a California-based non-profit organization, have spent $4 million to eradicate rats in the past decade and are now in the middle of a $3.5-million effort to regenerate the atoll's rainforest.

Palmyra

atoll

Palmyra could be a test case for restoring atolls and island ecosystems more broadly. Islands are some of the most imperilled, highly disturbed ecosystems in the world, and are home to roughly 20% of Earth's biodiversity — endemic species that evolved in isolation and exist nowhere else.

Remote WONDER

Located in the centre of the Pacific Ocean, Palmyra atoll is a ring of islets surrounded by coral reefs.

Most efforts to restore island ecosystems, including on Palmyra, have focused on removing invasive species; TNC has done that by eradicating rats. But since 2020, the organization has been replanting native plant species in the hope that a healthier forest could help to make Palmyra's coral reefs resilient to rising sea temperatures and increasingly acidic waters.

Native trees are part of a complex web that connects life on land and in the water. *Pisonia grandis*, or bird-catcher trees, are one of the native species that lure seabirds to the atoll. As bird guano collects around the gnarled roots and fills the soil with nutrients, soil microbes regulate their flow from the land to the sea, nourishing the reefs over time, says Kiers. The plan is to replant *P. grandis* to restore the island's health, but those efforts have so far met with limited success.

Toby Kiers describes how *Pisonia grandis* trees grow on the island. Credit: Charlie Cornwallis/Lund University

Show transcript

So you can see this is Pisonia.

Because one of the branches fell off.

But then it just started rooting up a whole new tree.

So it spreads vegetatively like this.

And all under this is the roots with the mycorrhizal fungi.

A similar story is playing out elsewhere: many restoration efforts around the world have faltered in their attempts to re-establish native species. Holly Jones, a restoration ecologist at Northern Illinois University in DeKalb, and her colleagues conducted a meta-analysis of 400 restoration projects worldwide and found that, although the studied ecosystems are recovering, they rarely do so completely[1]. And many projects do no better than letting ecosystems recover on their own.

That's where the fungi might come in, says Kiers, who is also executive director of the Society for the Protection of Underground Networks, an initiative to map fungal networks. They might be crucial for the health of *P. grandis* and a native shrub, *Heliotropium foertherianum*, both of which provide a dense canopy that birds prefer to nest in. Kiers is convinced that the mycorrhizae play a pivotal part in the cycling of the bird guano's nitrogen, phosphorous and potassium on land. With birds and mycorrhizae, "the trees are making their own soils", she says.

A red-footed booby (*Sula sula*) in a nest.

Credit: Charlie Cornwallis/Lund University

Working out how to improve success rates for ecological repair is a high priority, particularly during the United Nations Decade on Ecosystem Restoration. This initiative, running from 2021 to 2030, calls on nations to deliver on their commitments to restore a total of more than one billion hectares, an area the size of China. The cost of global terrestrial restoration — excluding marine ecosystems — is estimated to be at least $200 billion per year by 2030, according to a 2021 analysis by the UN Environment Programme.

Because many restoration efforts so far have had lacklustre results, a growing number of researchers think the soil microbiota could be a missing piece of the puzzle. Many think that adding soil microorganisms to degraded ecosystems could help native plant species to thrive and bring regions back to health. But some researchers, including Jones, urge caution because manipulating species, particularly on islands, has a long history of unintended consequences.

Those are some of the considerations that brought Kiers to Palmyra last year, to conduct the first studies on mycorrhizae in this remote atoll. Other researchers are watching closely to see what such studies have to offer, both for islands and ecosystems across the planet.

"If we can get restoration right on islands, we have this great capacity to have an outsized impact on reversing the world's biodiversity crisis," says Jones.

Island revival

Red-footed boobies fly alongside researchers as they head back to their field station. Credit: Toby Kiers/SPUN

Red-footed boobies fly alongside researchers as they head back to their field station. Credit: Toby Kiers/SPUN

Kiers and her colleagues spent 17 days on Palmyra in November 2022 to conduct the first survey of mycorrhizae and soil microbial DNA on 26 of the largest islets that have been cleared of Second World War munitions. The researchers recorded tree locations, dug through coral rubble to explore and collect tree roots, and took core samples of soil. The aim was to understand whether symbiotic fungi were present, and what role they might have in supporting the island's ecosystem. Eventually, these data could help TNC with its restoration efforts on Palmyra.

"Islands are a global conservation priority, but we lack data on what successful restoration looks like," says Alex Wegmann, director of science for TNC's Palmyra programme. The soil microbiome, he says, "is one of the last frontiers that we absolutely should explore to be able to effectively restore and protect these systems".

A satellite image of Palmyra atoll. Credit: NOAA

A satellite image of Palmyra atoll. Credit: NOAA

Palymra presents a big test for restoration efforts. When TNC purchased the atoll, about 45% of it was covered with coconut palms; *P. grandis* comprised just 13%. And coconut-palm seedlings exploded after the successful 2011 eradication of rats, which previously ate the seedlings as well as seabird eggs and chicks.

Since 2019, Palmyra land managers have painstakingly removed more than one million coconut palms. (The goal is to remove two million.) They have cleared palms from nearly 150 hectares and have so far replanted 18 hectares with native trees. For Wegmann, there is one overarching question: are the soils under the invasive palms able to host native *P. grandis* when the palms have been removed?

That's where mycorrhizae could be key. Of the four types of mycorrhiza, two are most common. One kind, called arbuscular mycorrhizal fungi, infiltrates the roots of an estimated 72% of most vascular plant species, notably agricultural crops. Another, known as ectomycorrhizae, forms a sheath surrounding the roots of roughly 2% of plant species[2], largely trees, including *Pisonia*. Both types of fungi produce threadlike filaments called hyphae that extend the roots' reach through the soil, improving the odds of securing often-scarce nutrients and water. These mats of hyphae not only serve as nutrient conduits, but also bind soil particles together to prevent erosion.

A stand of coconut palm trees killed as part of an eradication effort. Credit: Charlie Cornwallis/Lund University

A stand of coconut palm trees killed as part of an eradication effort. Credit: Charlie Cornwallis/Lund University

There are many tree-planting campaigns worldwide, but not all are successful, says Sofia Gomes, a postdoctoral researcher studying plant and soil microbial communities at Leiden University in the Netherlands. Even in abandoned crop fields, says Gomes, "it is surprisingly difficult to just go plant a tree". Trees that rely on mycorrhizae often need their specific below-ground partners to survive drought, disease or any other stresses when they arrive in a new environment, she says.

Studies in forestry and agriculture have repeatedly shown a drastic increase in tree or crop productivity when mycorrhizae are added, either by introducing commercial fungal spores or native soil to the planting mix. Mycorrhizal additions are also showing promise in efforts seeking to restore ecosystems. Reintroducing arbuscular mycorrhizae to abandoned agricultural fields in the US Midwest, for instance, consistently improved the success of tallgrass prairie restoration efforts[3]. A 2019 meta-analysis of 26 field-based studies[4] found that adding mycorrhizae to restored plant communities increased the number of species by 30%.

But land managers in the United States haven't yet taken advantage of these findings. A study published this year examined 130 management plans for forests and grasslands and found that only 8% mentioned mycorrhizal fungi, with only one plan noting that mycorrhizae could boost regeneration[5].

Beyond mycorrhizae and their partnership with plants, researchers think there are millions of other soil microorganisms that have crucial supporting roles in ecosystems. A 2022 analysis of 80 experiments showed that adding a small scoop of soil from intact ecosystems helps to boost plant biomass production by 64% on average, across ecosystems[6]. Researchers suspect the added microbes deserve the credit.

Kiers amid foliage below *Pisonia grandis* trees. The leaves are covered in guano from seabirds, which adds nutrients to the soil on the atoll. Credit: Charlie Cornwallis/Lund University

Kiers amid foliage below *Pisonia grandis* trees. The leaves are covered in guano from seabirds, which adds nutrients to the soil on the atoll. Credit: Charlie Cornwallis/Lund University

"Introducing a healthy soil microbiome from a native ecosystem tends to improve biodiversity even in the most degraded places, and speeds regeneration," says study co-author Thomas Crowther, an ecologist at the Swiss Federal Institute of Technology in Zurich who created Restor, an online platform for restoration project managers to efficiently share insights about successes and failures.

Focus on fungi

The part that mycorrhizae play is starting to come into focus in many ecosystems. On islands, it's clear that mycorrhizae shaped which plants initially colonized those areas. And on each of these isolated landmasses, the relationships between plants and fungi further influenced plant evolution[7].

But humans have moved plants around the world for centuries, unwittingly introducing root-hitchhiking mycorrhizae into island ecosystems. The pace of this activity has picked up in the past few decades.

"We don't know a lot about the native mycorrhizal communities in a lot of these islands," says Kabir Peay, who studies plant microbes at Stanford University in California and is extracting the soil DNA from Palmyra with his team. There are an estimated 40,000–50,000 fungal species that form mycorrhizal associations with plants[8], but it's a difficult number to estimate because so many habitats have yet to be sampled, he adds. As a result, there are very few situations where researchers can say whether a mycorrhizal fungus is introduced or native, says Nicole Hynson, a fungal ecologist at the University of Hawaii at Manoa.

Coconut palms invaded much of Palmyra after they were introduced in the nineteenth century. Eradication efforts have so far removed one million palms. Credit: Toby Kiers/SPUN

Coconut palms invaded much of Palmyra after they were introduced in the nineteenth century. Eradication efforts have so far removed one million palms. Credit: Toby Kiers/SPUN

Before arriving on Palmyra, Kiers found hints in the literature that *P. grandis*[9] and *H. foertherianum* on other islands might associate with mycorrhizae. And after Kiers sampled soils on Palmyra, she sent root material to Hynson, who confirmed that mycorrhizae are present on the roots of most of the *P. grandis* and *H. foertherianum* so far inspected. But the degree of colonization varies, and because the samples are blinded, Hynson doesn't yet know whether the roots with greater colonization are from native forest or restoration areas.

Kiers works in a laboratory on Palmyra to study microscopic fungal filaments in the roots of *Pisonia grandis* collected on the island.

Credit: Charlie Cornwallis/Lund University

A sample of roots from *Pisonia grandis* that were colonized by an ectomycorrhizal fungal symbiont.

In separate work in Hawaii, Hynson is exploring whether adding locally sourced fungi from the historical range of an endemic mint (*Phyllostegia kaalaensis*) that had been completely eradicated from the wild can help to restore populations. Previous attempts to reintroduce the mint had consistently failed because of disease. An initial study[10] found that after applying a slurry of fungi to the leaves of mint plants, they survived longer in the wild than ever previously recorded.

"Mycorrhizal fungi are like plant immune boosters," she says. Hynson is heartened by the results. But, she adds, "it's going to take more experimentation to find the right recipe for long-term success". It can take years, for example, to culture a mycorrhizal fungus.

There are some indications that these relationships could help restoration efforts on Palymra. In an unpublished pilot study started in 2019, researchers found that *Pisonia* cuttings planted in *Pisonia*-rich soil had a 50% survival rate, compared with only 37% survival for cuttings planted into soils collected from areas with no *Pisonia* trees. This suggests that soil gathered underneath a *Pisonia* grove contains beneficial microorganisms, says Wegmann.

Palmyra is home to 12 breeding species of seabird, which favour *Pisonia* forests. Sooty terns (*Onychoprion fuscata*) nest in this location. Credit: Charlie Cornwallis/Lund University

Palmyra is home to 12 breeding species of seabird, which favour *Pisonia* forests. Sooty terns (*Onychoprion fuscata*) nest in this location. Credit: Charlie Cornwallis/Lund University

With so few studies of plant-mycorrhizae dependency on islands, it is often unclear which, if any, below-ground species are crucial to plant success. Yet without interventions, some of the rarest, most endangered island plants could go extinct.

There are some hopeful signs. A 2020 study in Sulawesi in Indonesia found that mycorrhizae isolated from the roots of a tree in the same family as the endangered endemic tree, *Kalappia celebica*, significantly increased its growth on former gold-mining lands[11].

But some scientists are concerned about the unintended consequences of boosting mycorrhizal populations on islands. “We don’t know if the reintroduction of mycorrhizae will benefit native plants more than non-natives,” says James Bever, an ecological and evolutionary biologist at the University of Kansas in Lawrence. The fear is that adding mycorrhizae could give invasive plants an edge.

It’s a particular risk on islands because many have battled costly invasions of introduced species. Shallow-rooted miconia (*Miconia calvescens*), for example, is a non-native tree that was introduced in Hawaii; it accelerates erosion[12] and causes other types of damage estimated at $672 million annually.

Pisonia trees have a tangled network of roots that interact with mycorrhizal fungi in the soil. Credit: Toby Kiers/SPUN

Pisonia trees have a tangled network of roots that interact with mycorrhizal fungi in the soil. Credit: Toby Kiers/SPUN

There are hints that boosting mycorrhizae can backfire. This can be seen in initial work in the Galapagos Islands, conducted by Bever’s former graduate student, Jessica Duchicela, a restoration ecologist at the University of the Armed Forces in Sangolquí, Ecuador. Duchicela found that the non-native plant species benefited more from soil containing mycorrhizae than did native Galapagos plants[13].

“We want to restore the Galapagos to increase the density of natives, not the exotics,” she says. Researchers need more information about what is in the soil and how the plants are responding before any attempts to alter the soil microbiota to restore endemic plants, she cautions.

“We have to ask these questions if we want to get it right,” says Kiers.

But it’s difficult to conduct some of these studies, especially in areas such as the Galapagos and Hawaii, which have strict protocols to protect native flora. Duchicela has spent the past few years trying to get permits to take soils from the Galapagos Islands to her laboratory to continue her mycorrhizae research.

Crab transport

While studying the *Pisonia* forests on Palmyra, Kiers and her colleagues marvelled at the clusters of crabs churning the soil by digging holes in and among the roots of trees — potentially spreading the mycorrhizal spores. She eagerly awaits the results of DNA sequencing on soil from crab holes to see whether there is evidence of the crustaceans moving fungi.

In her expedition last year, Kiers explored life both above and below the waves. As she was wading through shallow water one day, she shrieked after a curious reef shark bumped into her. The prowling sharks, captured in a video taken during the trip, are a sign of vibrant reefs, nourished in part by the one million birds that call Palmyra home — including massive colonies of sooty terns (*Onychoprion fuscatus*) and red-footed boobies (*Sula sula*). “We could hardly hear ourselves talk because the birds were so loud,” she says of the islets dominated by *P. grandis.* It was a sharp contrast to the relative quiet of the islets covered by coconut palms. Of Palmyra’s 12 breeding seabird species, only 2 will regularly nest in coconut palms, says Wegmann.

Reef sharks swam around the researchers as they walked through shallow water between different islands. (This video has no sound.) Credit: Toby Kiers/SPUN

Reef sharks swam around the researchers as they walked through shallow water between different islands. (This video has no sound.) Credit: Toby Kiers/SPUN

In the past five years, scientists have shown that islands with more seabirds — and therefore more nutrients — have more-resilient coral reefs. A 2018 study compared islands rife with seabirds to rat-infested islands that contained fewer seabirds. The coral reefs and fish that live around them grew faster near islands with more seabird nutrients[14]. “Before this paper, no one had connected the potential benefits of restoring natural nutrient flows to reefs,” says Casey Benkwitt, a research fellow at Lancaster University, UK, who is tracing how seabird nutrients reach reefs.

Although seabirds are the source of nutrients, Kiers thinks the soil microorganisms are the subterranean transport system that keeps nutrients flowing steadily on and off the island. In areas colonized by *P. grandis* and ectomycorrhizae, the soil is moist brown peat, entirely composed of organic matter, the kind that smelled like vegetables to Kiers. By contrast, the islets covered in coconuts were often just pure sand, unable to hold onto nutrients, she says.

Kiers describes how *Pisonia* trees and mycorrhizal fungi help to form soil on Palmyra. Credit: Charlie Cornwallis/Lund University

Show transcript

That is what is unexpected.

Is that this Pisonia tree seems like it's actually making its own soil.

It has these particular holes that collect organic matter and it becomes thick rich soil which then the roots grow back into and the fungi take up the nutrients and bring it up to the plant.

Peay hopes researchers can race to document the below-ground biodiversity on islands — before further changes above ground make it even more difficult to disentangle native soil microorganisms from ones that humans introduced. “It’s critical to get a baseline understanding of which species are present and probably native,” says Peay, “because people are moving things around more and more.”

“We might be erasing the soil’s bio-geographic signature, which could have consequences for restoration,” he adds.

If, Kiers stresses, they find that the ectomycorrhizae were lost as palms took over the islands, it might be necessary to transplant soil from Palmyra’s remaining *Pisonia* forests to reintroduce the fungus onto islets that have been cleared of palms. “There is no other way *P. grandis* will be established without its fungal symbiont,” she says.

If trees can be restored to protect the corals, healthy reefs will return the favour. As sea levels rise, resilient coral reefs will provide the natural barriers that buffer waves and allow islands to persist, says Kiers. “Survival of the coral reef is fundamental to the survival of the atoll.”

Credit: Toby Kiers/SPUN

Credit: Toby Kiers/SPUN

- **Author:** Virginia Gewin
- **Media editor:** Agnese Abrusci
- **Subeditor:** Anne Haggart
- **Art editor:** Chris Ryan
- **Editor:** Richard Monastersky

Photos, video and drone footage from Palmyra: SPUN and The Nature Conservancy

References

1. Jones, H. P. et *al*. *Proc. R. Soc. B* **285**, 20172577 (2018).
2. Cope, K. R. *et al*. *Plant Cell* **31**, 2386–2410 (2019).
3. Koziol, L., McKenna, T. P. & Bever, J. D. *Plants* **12**, 1184 (2023).
4. Neuenkamp, L., Prober, S. M., Price, J. N., Zobel, M. & Standish, R. J. *Fungal Ecol*. **40**, 140–149 (2019).
5. Markovchick, L. M. *et al*. *Restor. Ecol*. **31**, e13866 (2023).
6. Averill, C. *et al*. *Nature Microbiol*. **7**, 1717–1725 (2022).
7. Delavaux, C. S. *et al*. *Nature Ecol. Evol*. **3**, 424–429 (2019).
8. van der Heijden, M. G. A., Martin, F. M., Selosse, M.-A. & Sanders, I. R. *N. Phytol*. **205**, 1406–1423 (2015).
9. Suvi, T. *et al*. *Mycologia* **102**, 522–533 (2010).
10. Zahn, G. & Amend, A. S. *PeerJ* **5**, e4020 (2017).
11. Arif, A., Husna, Tuheteru, F. D. & Adiansyah, A. *IOP Conf. Ser. Earth Environ. Sci*. **800**, 012039 (2021).
12. Burnett, K., Kaiser, B. & Roumasset, J. A. *J. Forest Econ*. **13**, 151–167 (2007).
13. Duchicela, J., Bever, J. D. & Schultz, P. A. *Plants* **9**, 74 (2020).
14. Graham, N. A. J. *et al*. *Nature* **559**, 250–253 (2018).

SPRINGER NATURE
Springer Nature

- Privacy Policy
- Use of cookies
- Legal notice
- Terms & Conditions
- Accessibility statement

© 2023 Springer Nature Limited. All rights reserved.
TopBuilt with **Shorthand**

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01932-y

| Section menu | Main menu |

Books & Arts

- **Is ‘speciesism’ as bad as racism or sexism?** [19 Jun 2023]
  • We are all complicit in a global farming industry that puts profit before animal welfare — but establishing what moral principles we should be applying isn’t easy, an update of a classic book shows.

- BOOK REVIEW
- 19 June 2023

Is ‘speciesism’ as bad as racism or sexism?

We are all complicit in a global farming industry that puts profit before animal welfare — but establishing what moral principles we should be applying isn’t easy, an update of a classic book shows.

- Jonathan Birch



Broiler chickens have been selectively bred to have voracious appetites and to grow rapidly.Credit: Simon Belcher/Alamy

Animal Liberation Now *Peter Singer* Bodley Head (2023)

My three-year-old son adores animals. He has a farm in his room in which small plastic pigs live in pleasant pastoral surroundings. How can I break it to him? Modern farming is not like this. Pigs and chickens in intensive-farming systems endure conditions no toy manufacturer would dare to reproduce.

Chickens, up to 50,000 to a shed, have been bred over the past 100 or so years to grow at four times their natural rate and to reach three times their natural weight or more, gaining weight so fast that their legs and hearts struggle to support them in the last weeks of their 40-day lives (M. J. Zuidhof *et al. Poultry Sci.* **93**, 2970–2982; 2014). Sows, bred to reach enormous sizes, are confined in farrowing crates to stop them accidentally crushing their piglets. Most of those piglets will have their tails docked: otherwise, in sheds with scant other stimulation, they bite at and maim each other's tails. Five to six months later, most of them are gassed with carbon dioxide, an extremely unpleasant experience that can take a minute or more.

We know as much as we do about the welfare implications of intensive farming because of the work of scientists and campaigners. Many of them were inspired to enter these areas by Peter Singer's book *Animal Liberation*, first published in 1975. It was one of the first books to document the suffering caused to animals in detail, and present it together with strident ethical arguments calling for action. Almost 50 years later, Singer has produced a massively revised and updated edition, *Animal Liberation Now*. The new material is not all bad news. The book showcases signature achievements of the animal-welfare movement, such as the European Union's commitment to end all cage farming by 2027 and California's recent ban on veal crates, gestation crates and battery cages. But it is mostly bad news. Many problems have worsened.

In a long chapter on intensive farming, Singer lets loose a barrage of unconscionable facts. In the case of broiler chickens, selective breeding aimed at optimizing profit rather than welfare has gone too far. It is a global

regulatory failure, and to carry on using superfast-growing breeds now that the consequences for animal welfare are so well known is inexcusable. Dutch supermarkets have committed to phasing out their use by the end of 2023; one UK food retailer, Marks & Spencer, did so in 2022. Other producers and retailers should take note.

In this part of the book, ethical commentary takes a back seat to the dispassionate communication of facts. That is a wise choice, because the facts speak for themselves.

Facing the facts

But what can we, as individuals, do about it? As Singer noted in 1975, the most obvious answer is to boycott the products of intensive farming through veganism, vegetarianism or ‘conscientious omnivorism’ (eating only meat from higher-welfare extensive farms). It is sobering, then, to see Singer reflect that “my call for a boycott of meat has been a dismal failure”. In a few places, notably Sweden and Germany, meat consumption has begun to fall in recent years; but in most, demand has only risen in the decades since *Animal Liberation*, and it is projected to grow further. Singer is acutely aware of the objection that, from a producer’s point of view, a consumer’s behaviour is just noise in the data against a clear trend of rising demand. An individual changing their habits does not send a signal that would force companies to cut supply. So, what is the point?



[What humanity should eat to stay healthy and save the planet]

Singer has always struggled to reply to this convincingly because he is committed to a way of thinking about ethics that centres on individual actions. I think the right reply is to shift the focus from individual to collective actions. Pushing for systemic change is something we must do together. Our individual contributions will be negligible, but our coordinated action has the power to shift industry norms. Veganism, vegetarianism and conscientious omnivorism, to work at all, must operate as organized social movements, not as isolated individual choices. The goals of such actions should include policy changes, including a reduction in the government subsidies that flow to intensive animal agriculture, not just dietary adjustments.

The climate connection

Later in the book, Singer makes valuable connections between animal welfare and the fight against climate change, an issue that was (by his own admission) not on his radar in 1975. He imagines a critic who retorts: if I eat tofu instead of meat, will I not be contributing to the destruction of the Amazon rainforest? The answer is no: 77% of the world's soya is fed to animals, with only 7% used for human consumption.

Shifting away from meat is highly recommended on environmental grounds. Yet there are genuine points of conflict, which Singer underplays. Environmental campaigners tend to see reducing beef consumption as their top priority, because the emissions associated with beef production are so great. But they sometimes recommend replacing beef with chicken, potentially further increasing demand for one of the lowest-welfare kinds of farming.

Moreover, there are some rhetorical and philosophical choices that will always limit the book's appeal. In both editions, Singer rests his arguments on the rejection of 'speciesism': a type of discrimination that Singer likens to sexism or racism, but on grounds of species. For Singer, we might differ in how we treat an animal on the grounds of cognitive capacity, but never just because it is a different species. For example, he argues, if you would be willing to kill a pig and harvest its organs to benefit people (if they were human-compatible), you should be willing to do the same to a human infant born without parts of its brain and skull, which has less mental capacity than the pig.



An exhibition at Miniatur Wunderland in Hamburg, Germany, shows the intensive farming of pigs.Credit: dpa picture alliance/Alamy

Singer is a committed utilitarian, advocating solutions that maximize overall well-being: the greatest happiness for the highest number. He thus favours taking organs from both the pig and the infant if the benefits are large enough, relative to any suffering caused. This is a Singer litmus test: some readers will welcome the comparison, most will seethe. Disability-rights campaigners have long found this style of argument exasperating and offensive. The tragedy of Singer's approach is that two progressive movements that should be allied end up seemingly opposed.

Personally, I find this way of framing the issues unhelpful, and the new edition is a missed opportunity to take a different tack. The analogy with racism and sexism is too weak and fragile a foundation on which to rest arguments for respecting other animals. Ultimately, it is no bad thing to have a medical profession concerned only with humans that adheres to a norm of 'first, do no harm'. Public trust in medicine relies on this. Even in a

utilitarian framework, we can find good reasons for the medical and veterinary professions to have different norms. More fundamentally, I think the focus on speciesism misdiagnoses the source of the problem. Our maltreatment of other animals stems not from speciesism as such, but from failing to respect other animals as sentient beings with lives of their own.

Animal welfare in research

A second questionable choice is to put animal experimentation in the same bracket as intensive farming, portraying it as another manifestation of egregious speciesism. In the case of animal research, the harms and benefits are often balanced much more finely, and Singer's trenchant style seems less appropriate. Although he is not an absolutist abolitionist — proposing that the benefits can, in principle and in exceedingly rare circumstances, outweigh the harms — the vast majority of animal research is in his crosshairs.



Eat more fish: when switching to seafood helps — and when it doesn't

Yet an awkward truth is that protecting animal welfare requires an understanding of animal minds and behaviour, and the understanding we have rests substantially on animal research. Singer cites, as an example of scientists crossing ethical boundaries, a study in which prairie voles were

imprisoned in small tubes in front of others, inducing “anxiety-like states”. In a recent lecture on “animal emotions and animal rights law” at the University of Cambridge, UK, the primatologist Frans de Waal used the same paper to argue that rodents have empathy, and so have more-complex welfare needs than previously thought (go.nature.com/3qghauc). Was the knowledge worth the harm? Institutions need robust oversight to make sure such questions are asked, but the scales are not as lopsided as Singer suggests.

The book is a reminder that plain, well-sourced facts, starkly presented, often speak louder than philosophical arguments. I recently took my son to Miniatur Wunderland, a vast model railway in Hamburg, Germany. I was surprised to find, in one of the landscapes surrounding the tracks, rows of little sheds in which tiny plastic pigs were being intensively farmed. The model makers have had, it turns out, a long-standing interest in animal welfare. Apparently, they had tried some crude shock tactics, but after facing a backlash, they switched to a more subtle approach: depicting reality in painstaking detail, without comment. It was a powerful exhibit. Perhaps that is how my son’s journey starts. *Animal Liberation* and *Animal Liberation Now* owe their undeniable power to the same basic source: to know the facts is to see at once that change is needed.

Nature **618**, 667-668 (2023)

doi: https://doi.org/10.1038/d41586-023-01977-z

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01977-z

Opinion

- **Gordon Moore (1929–2023)** [16 Jun 2023]
  • Microchip entrepreneur and architect of Moore's Law.
- **Current conservation policies risk accelerating biodiversity loss** [21 Jun 2023]
  • Three approaches that aim to cut the harms of agriculture — land sharing, rewilding and organic farming — risk driving up food imports and causing environmental damage overseas. An alternative approach is both effective and cheaper.

- OBITUARY
- 16 June 2023

Gordon Moore (1929–2023)

Microchip entrepreneur and architect of Moore's Law.

- Christophe Lécuyer



Credit: Roger Ressmeyer/Corbis/VCG via Getty

Gordon Moore played a pivotal part in the development of the semiconductor technology and industry in the United States, co-founding both Fairchild Semiconductor and Intel Corporation. He made the observation about the development of computing power now known as Moore's Law, and established the technology road map for semiconductors that drove innovation in microelectronics, in accordance with the law, for nearly one-quarter of a century. He died on 24 March, aged 94.

Moore grew up in Pescadero, a small farming town on the California coast, south of San Francisco. At a young age, he took an avid interest in chemistry and experimented with explosives in his home laboratory. He obtained a bachelor's degree in chemistry at the University of California, Berkeley, followed by a doctorate in physical chemistry at the California Institute of Technology (Caltech) in Pasadena. His life changed in 1956, when William Shockley, a physics Nobel laureate and co-inventor of the transistor, recruited him as a research chemist at the Shockley Semiconductor Laboratory he had recently founded in what became Silicon Valley. There, Moore learned the complex and highly experimental techniques required to produce diodes and transistors made of silicon.

In 1957, Moore and seven other scientists and engineers on the Shockley staff rebelled against their abusive and mercurial boss, establishing the Fairchild Semiconductor Corporation to specialize in high-quality silicon devices. Moore quickly became the head of the firm's research laboratory. Under his leadership, the lab made two major innovations, the integrated circuit and the 'planar' transistor manufacturing process that allowed the production of reliable devices with excellent electrical characteristics. Moore gave his researchers a lot of freedom. For much of the 1960s, his lab remained the most productive in the US microelectronics industry.



Moore’s law is not just for computers

In 1965, Moore published his now-famous article, ‘Cramming more components onto integrated circuits’, in *Electronics*, a trade magazine. He argued that integrated circuits would rapidly grow in complexity, driving down the cost of electronic functions. ‘Moore’s plot’ showed that the number of components per microchip had doubled each year since the development of the planar transistor in 1959. He predicted that this trend would continue for ten more years. In his view, competitiveness and continued profitability in the microelectronics industry depended on it.

Moore used his plot to identify a new technical and commercial opportunity — semiconductor memory chips. This led him and Robert Noyce to leave Fairchild and found NM Electronics in 1968, the company that later became Intel. Under Moore’s direction, Intel’s engineers developed a process that used a new material, polysilicon, to form the gates of transistors that control whether they are switched on and off. This revolutionary process enabled the corporation to introduce microchips with the most transistors to the market — memory chips first and microprocessors later. Moore and his associates at Intel and Caltech also heavily promoted Moore’s plot in the semiconductor community to boost investment in the continued complexification of integrated circuits. As a result, Moore’s plot became increasingly known as Moore’s Law in the late 1970s.

Moore became Intel’s president in 1975 and chief executive in 1979, holding this position until 1987. A quiet but effective leader, he was attentive to the people around him and thought nothing of spending several hours setting up a junior staff member’s computer. Intel grew rapidly in the late 1970s, but was increasingly outpaced by Japanese firms with much lower manufacturing costs. By 1985, the company was on the brink of bankruptcy. To save it, Moore concentrated its technical capabilities on microprocessors, reforming manufacturing and stepping up the rate of technological change. He also stopped allowing other firms to produce Intel’s microprocessors, ensuring that the company had a monopoly on its own designs. This, and the rapid growth of the personal-computer market, enabled Intel to become the world’s largest chipmaker by 1992.



In AI, is bigger always better?

Moore was by then chairman of Intel’s board. To further solidify the position of Intel and other US firms in the global market and overcome obstacles to exponential complexity growth, he established the National Technology Roadmap for Semiconductors (NTRS). Gathering experts from industry, government and academia, the NTRS specified the main metrics for microchip complexity over the next 15 years for the entire US microelectronics industry, and translated them into a set of development needs for manufacturing processes. The road map set the funding priorities of the US federal government for semiconductors and oriented the research

and development activities of chipmakers, equipment suppliers, academic institutions and even the National Laboratories. In 1998, the NTRS became international, guiding the development of semiconductor technology at the global scale until it folded in 2016.

In the latter part of his career, Moore focused on philanthropic pursuits. With his wife Betty, he started the Gordon and Betty Moore Foundation to improve patient care and finance research in the sciences. The foundation currently makes grants of more than US$400 million per year in health care, science and the environment. Perhaps in atonement for Intel's poor environmental record during the 1970s, Moore also became a major patron of Conservation International, a pro-business environmental organization protecting ecosystems in more than 70 countries.

Nature **618**, 669 (2023)

doi: https://doi.org/10.1038/d41586-023-01978-y

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01978-y

- COMMENT
- 21 June 2023

Current conservation policies risk accelerating biodiversity loss

Three approaches that aim to cut the harms of agriculture — land sharing, rewilding and organic farming — risk driving up food imports and causing environmental damage overseas. An alternative approach is both effective and cheaper.

- Ian Bateman &
- Andrew Balmford



Tea pickers in Sri Lanka remove weeds from an organic tea plantation. Credit: Ishara S. Kodikara/AFP via Getty

With agriculture the main driver of the habitat loss and degradation that underpin the global biodiversity crisis[1], governments worldwide have implemented policies to lessen farming's impact on the environment. Meanwhile, landowners, organizations interested in the financing of biodiversity conservation and certain non-governmental groups, including conservation bodies, have been pushing for land-use changes that benefit nature.

However, numerous studies show that some of today's most popular conservation policies are doing little to help those species most affected by farming. What's more, by reducing how much food is produced per unit area (yield), they are driving up food imports and thereby having an impact on wildlife overseas.

One of us (I.B.) has advised seven UK secretaries of state for the environment over the past decade; the other (A.B.) has, for two decades, led empirical work investigating how to reconcile food production with biodiversity conservation. In our view, there are many reasons for the disconnect between the science and policy.

Part of the problem is that, especially in Europe, the owners of the biggest, and often richest, farms stand to gain the most from current policies[2]. Thousands of influential individuals are lobbying to maintain the status quo in agricultural policy.



EU climate plan sacrifices carbon storage and biodiversity for bioenergy

A more fundamental and massively under-recognized problem is that both government policy and much academic debate have focused too narrowly on the local effects of a given approach, rather than on its overall (often long-distance) impacts. Indeed, this tendency to ignore downstream consequences — even as much better tools and data become available to track and quantify such impacts — is causing significant problems across a range of conservation and climate policies.

It doesn't have to be this way. Modelling approaches are now being developed so that the information that is already available can significantly improve decision-making around agriculture and the environment. Using the wealth of evidence from research to guide agricultural policy could better reconcile conservation with people's need for food. It could also pave the way for the evidence-based decision-making that is urgently needed across a broad sweep of environmental challenges.

In vogue

In response to the biodiversity crisis, the European Union, the United Kingdom, Japan, Mexico and other regions and countries are increasingly

devoting resources to what seem to be environmentally friendlier ways to use land.

Since it was created in the 1960s, the Common Agricultural Policy (CAP) has been the EU's most expensive policy. More than one-fifth of the CAP budget — currently set at €56 billion (US$60 billion) a year — is available for 'environmental improvement', and most of that is funnelled into 'land-sharing' schemes.



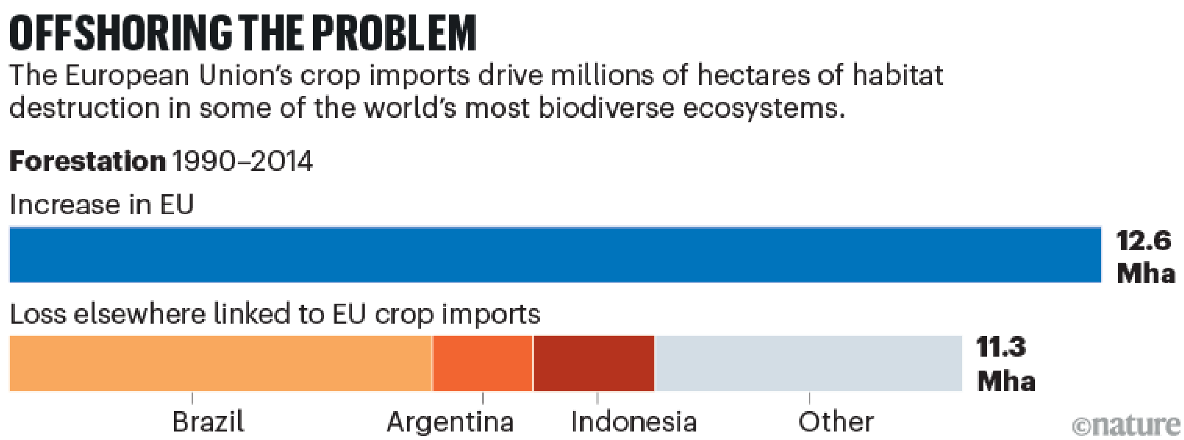
A wildflower margin next to a harvested Dutch wheat field — an example of land sharing.Credit: Getty

Land sharing uses a variety of approaches to increase biodiversity in farmland. Interventions include reducing the use of pesticides and fertilizers, adopting more-diverse cropping regimes and creating small-scale habitats such as unsprayed field margins and small patches of woodland.

Land sharing increases the populations of relatively common animals and plants, such as skylarks, field poppies and the more widespread butterflies. And highly targeted interventions can help some vulnerable species. But in

the main, land sharing does little for those most specialized or threatened species that need large stretches of contiguous non-farmed habitat, such as the many birds, invertebrates, plants and fungi dependent on old-growth forest. In fact, farmland biodiversity has continued to decline under land-sharing policies, with the UK population of corn buntings (*Emberiza calandra*), for example, falling by 83% since the late 1960s[3].

What's more, taking land out of agriculture without lowering food demand or raising yields elsewhere in a country increases the need for imports, which means more harm to biodiversity and natural habitats farther away[4]. Indeed, the EU's crop imports in the 25 years up to 2014 generated more than 11 million hectares of habitat destruction[5] — an area larger than Cuba — in some of the world's most biodiverse ecosystems, including those in Brazil and Indonesia (see 'Offshoring the problem'). In 2020, even the EU's official auditors declared the CAP a failure in terms of its environmental policies (see go.nature.com/45sasew).



Source: Ref. 5

Besides land sharing, two approaches have been gaining popularity in recent years, but each brings similar problems.

Some conservation groups and landowners have increasingly advocated for rewilding, in which large, contiguous areas of land are taken out of farming. Rewilding can benefit species that are locally vulnerable or endangered. For example, efforts to rewild 400 hectares at Ken Hill in eastern England have created a refuge for beavers, which have been extinct in the United Kingdom

since the sixteenth century. Such rewilding initiatives are obviously of value to national biodiversity. However, assessments of the benefits rarely consider offshore damage. As with land sharing, unless people change their diets or eat less, or yields are increased in areas that are still farmed, the removal of land from productive agriculture will increase the demand for food imports, and so damage biodiversity elsewhere[6].



Europe's Green Deal offshores environmental damage to other nations

Organic farming has been around for longer. In the past few years, both the EU and Japan have committed to converting one-quarter of farmland to organic production by 2030[5] and 2050 (see go.nature.com/43qycet), respectively. Although some farmland species are likely to benefit from the removal of manufactured fertilizers and modern pesticides, the approach will do little to help those that require contiguous natural habitats. Furthermore, organic production drastically lowers yields. Sri Lanka's recent food crisis has been attributed in large measure to the government's (now abandoned) attempt to convert the country to organic farming (see go.nature.com/3p2kgfq). And estimates suggest that a wholesale switch to organic farming across England and Wales would cut food-calorie output by 40%[7]. Again, this would lead to greater demand for food imports, and so increase pressure on production and hence on biodiversity around the world.

What the science says

Fortunately, another approach could bring substantial benefit to both local and global biodiversity (see ‘Helping or harming nature?’). Land sparing involves lumping habitat patches together into larger blocks, alongside the adoption of lower-impact ways to boost yields in the areas that are still farmed. Together, these two actions can make space for better habitat protection locally without displacing production overseas.

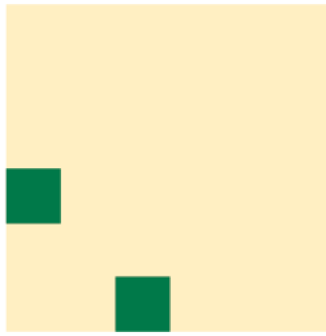
HELPING OR HARMING NATURE?

Land sharing is a popular way of altering land use to benefit nature, but it worsens the global biodiversity crisis by reducing the productivity, or yield, of farmland and driving up food imports. Another strategy, land sparing, could work.

Land use

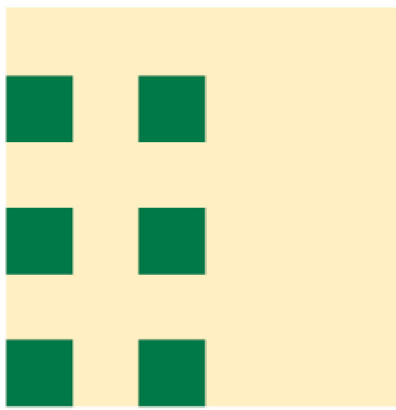
- Conservation areas
- Average-yield agriculture
- High-yield agriculture

Conventional farming



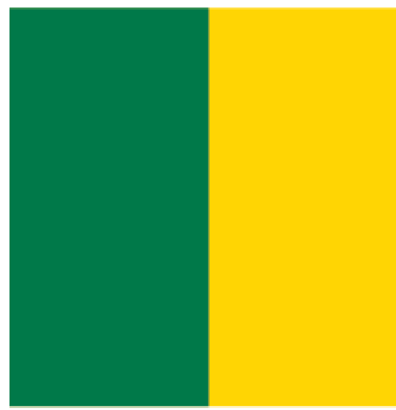
Has reduced large areas of contiguous wild habitat required by many specialist species. Common non-specialists have |also declined.

Land sharing



Increases conservation habitat but uses relatively small, fragmented areas, so aids common but not many threatened species. Reduced food production increases reliance on imports, worsening biodiversity loss overseas.

Land sparing



This spares the larger contiguous areas needed by threatened species in the region, and potentially reduces imports.

Impact compared with conventional farming

- Better
- Same
- Worse

Land sharing:

National population
- Threatened species (Same)
- Common species (Better)

Global population
- Threatened species (Worse)

Food
- Domestic production (Worse)

Land sparing:

National population
- Threatened species (Better)
- Common species (Worse)

Global population
- Threatened species (Same/Better)

Food
- Domestic production (Same/Better)

©nature

Source: data from multiple studies; see ref. 9 for a review.

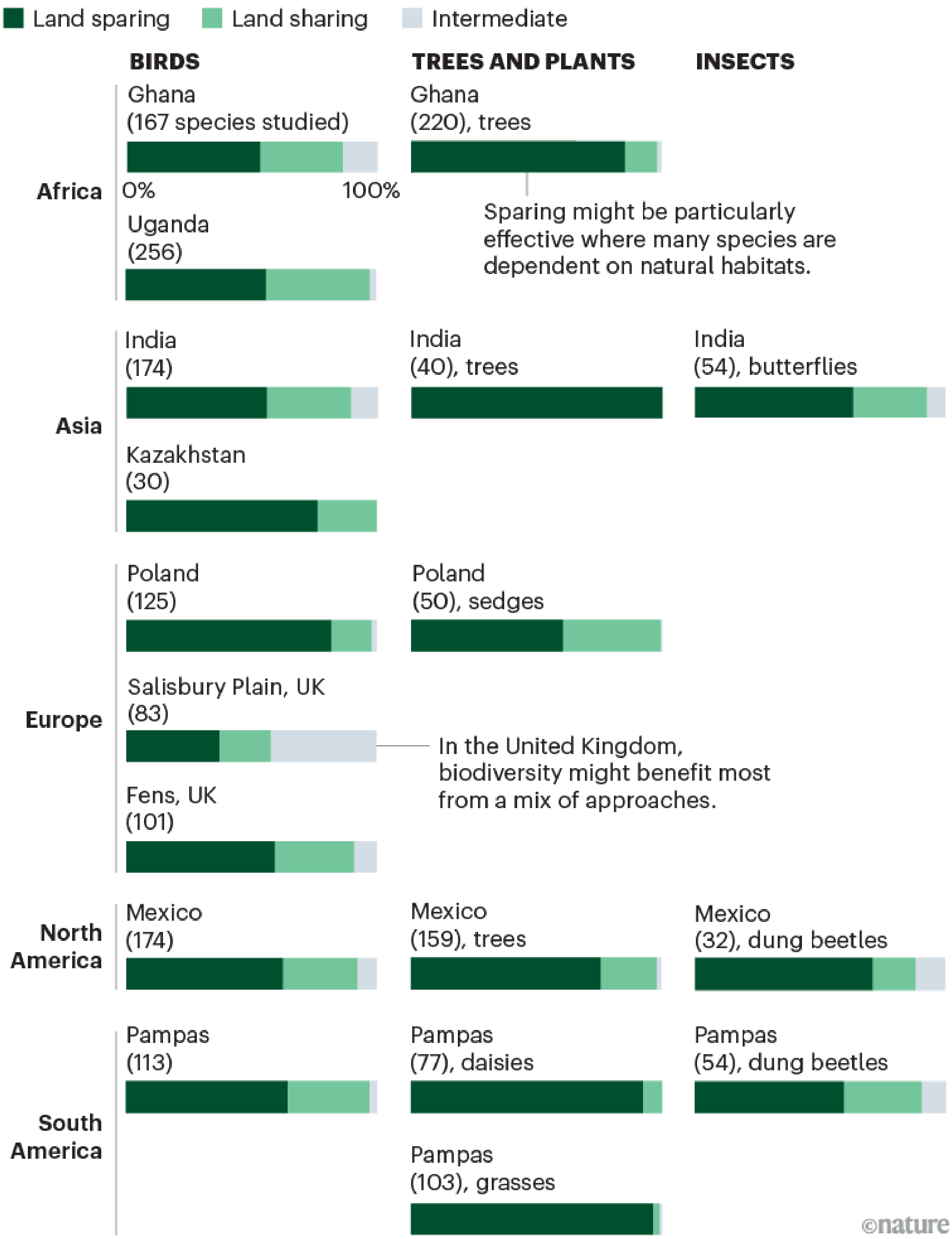
Choosing which areas to put aside for nature requires an understanding of the consequences of land-use change — for food production, but also for greenhouse-gas fluxes, hydrological regimes, access to recreation, the spread of pollutants and so on. But, in relation to biodiversity, larger habitat blocks — which are less affected by the drier, often windier and more variable conditions at the margins — can, for their size, hold larger populations of those species that favour more-natural habitats. The greater diversity of environments that arise in larger areas also supports a greater diversity of specialist species[8].

For areas that are still farmed, an array of techniques can help producers to raise crop and livestock yields sustainably. Options include providing animals in extensive grazing systems with greater access to improved pasture, water sources and modern veterinary care; using genomic screening and gene editing to accelerate animal and crop breeding; growing high-value crops such as salad vegetables and herbs in trays that are stacked vertically; using native plants to redistribute pests away from crops; and using 'recirculating' aquaculture systems to produce high-value products such as king prawns.

Over the past decade or so, field studies in India, Ghana, Uganda, Kazakhstan, Mexico, Colombia, Brazil and Uruguay, as well as in Poland and the United Kingdom, have all concluded[9] that (for the same overall food output), high-yield farming combined with land sparing results in larger populations of most wild species than does land sharing (see 'Winners and losers'). These findings, across more than 2,000 species of bird, plant and insect, are especially pronounced for those species with narrow geographical distributions, which make them particularly vulnerable.

WINNERS AND LOSERS

Numerous studies have compared how species would fare under different farming strategies. The bar charts below show the proportion of species assessed whose populations would be maximized under three types of land use.



Source: data from multiple studies; see ref. 9 for a review.

Last year, a study that surveyed UK farmers to establish what actions they would take, for what payment, found that delivering the same biodiversity outcomes for birds through land sparing would cost the taxpayer just 48% of the cost under land sharing; the impact on domestic food production would also be 21% lower under land sparing[10]. Thus, for the same budget, sparing seems to provide much greater biodiversity gains than does sharing.

Other research has shown that, in comparison to land sharing, a land-sparing approach can deliver significantly greater co-benefits, such as the removal and storage of greenhouse gases and the provision of recreational areas[11]. And calculations for the United Kingdom and Poland show that blended approaches, which combine spared land with shared farmland and high-productivity agricultural land, do even better in these countries than does pure sparing, and greatly outperform both current farming systems and pure sharing approaches[9].

Because sparing increases yields on the land still being farmed (and this is more easily achieved wherever there are big gaps between current and potential yields), these approaches can help to address food-security concerns[10]. Also, the need for both agricultural innovation and, in many areas, habitat restoration means that land sparing need not adversely affect rural employment[12].

Of course, yield increases do not inevitably lead to more land being available for conservation. Critics of land sparing point out that gains in yield could simply lead to rebound effects, with less land being taken out of farming than expected, or even to more land being converted to farmland because of the promise of greater profits[13].



Aquaculture techniques can raise yields sustainably, such as at this prawn farm in La Cruz, Mexico.Credit: Susana Gonzalez/Bloomberg via Getty

The evidence suggests, however, that although individual food producers generally use yield-intensifying practices to boost their incomes, overall land use tends to decrease[13]. These benefits could be increased by policies and subsidies crafted to dampen rebound effects; farmers could be given support for innovation in exchange for reducing the area under cultivation. A reassessment of data for birds and trees in Ghana and India shows that sparing would still outperform sharing even when policies to limit rebound effects are not put in place[9].

A matter of focus

So, given that land sparing could benefit more biodiversity at lower cost than can other strategies, and deliver an array of co-benefits, why is it not the dominant approach today?

The influence of the 'big farm' lobby in maintaining the status quo in agricultural policy is one widely cited reason[14]. The chief approach to allocating subsidies — using flat-rate payments per hectare of shared land — disproportionately benefits the largest (and often richest) farms. As a result, in the United Kingdom, 12% of farms take 50% of all taxpayer subsidies, whereas half of all farms share just 10% of those subsidies[2] (see 'The devil is in the detail').

The devil is in the detail

Subsidies are used to persuade landowners to make changes to benefit nature, but subsidy design determines how effective they are.

Flat-rate subsidies pay farmers a set amount per hectare for conservation. This common approach channels the majority of subsidies towards the largest, and often richest, farms. Such schemes also fail to incentivize farmers to do more than the minimum stipulated in policy documents and often penalize those who go further. For example, if UK farmers plant trees on their land (and do not fell them at least every ten years), that land is removed from the tax exemptions accorded to 'productive' farmland (see go.nature.com/468wrzb).

A better approach would be to pay rewards not for the amount of land farmers devote to 'nature improvement', but for the expected outcomes. For this, farmers would be free to propose actions to address a specific environmental problem and to state what payments they would accept in return. Modelling is then used to predict benefits. By comparing these expected outcomes with the costs required by each farmer, decision makers can choose those farms and actions that deliver the best value for money[19].

In our view, however, a more fundamental and much less recognized problem confounds the application of scientific research to environmental policy — and not just in relation to agriculture.

The 'focusing illusion'[15], proposed by the Nobel-prizewinning psychologist and economist Daniel Kahneman, is the psychological phenomenon that focusing on one effect of a change tends to diminish our perception of all the

other possible effects of that change. The literature is replete with studies of the effects of a change in terms of a single (often local) measure: biodiversity or carbon storage, nitrogen pollution or flooding, food production or recreation. Fewer assessments exist of multiple outcomes or of system-wide impacts.

Historically, part of the challenge has been a lack of data and understanding. Even studies considering the plural effects of a change in how land is used have often been locally or nationally focused, largely because the modelling work linking the change to broader economic and environmental effects hasn't been available[16].

Global-trade modelling, however, is now enabling researchers to obtain a much fuller picture of the economic and environmental effects of both policy interventions and business investments[17].



Food systems: seven priorities to end hunger and protect the planet

Over the past five years or so, there has also been more research aimed at designing tools that allow policymakers and other stakeholders to understand the wider consequences of a change in land use. As an example, one of us (I.B.) is involved in a project to examine the full effects of the UK government's decision, in 2020, to fund substantial increases in national woodland cover to remove greenhouse gases[18]. The Natural Environment

Valuation Online tool (www.leep.exeter.ac.uk/nevo), which will be used in this project, combines information from multiple disciplines to show decision makers how such a change in the way land is used will help to satisfy England and Wales's net-zero-emissions commitments, benefit biodiversity, improve access to recreation and so on. The tool also shows the impact of changes in land use on domestic food production, which can then be linked to changes in land use and biodiversity globally.

The goal of research on system-wide impacts should not be to obtain ever more detailed sources of information about all the possible effects of a proposed policy change. Rather, analyses should be extended to the point at which the costs of collecting and analysing more data begin to exceed the benefits of more-informed decision-making. Such interdisciplinary studies and approaches that focus on the needs of decision makers must become the norm.

The stakes are too high for policymakers to continue to ignore the promise of land sparing when so much research demonstrates that it is a much more effective approach than many of the strategies being deployed. This issue has become even more urgent since last December, with the adoption of the Convention on Biological Diversity's goal of protecting 30% of the planet's land and oceans by 2030. Exactly how this 30% will be put aside (as large contiguous natural habitats or as a multitude of fragments), and how the world's growing demand for food and other goods will be met from the unprotected remainder of Earth's surface, will in large part determine the biodiversity consequences of this ambitious commitment.

Yet the story about land sparing carries an even broader message: unless researchers and policymakers assess the overall, global effects of interventions aimed at addressing biodiversity loss, climate change and environmental degradation, poor decisions that are unsupported by the data will at best under-deliver, and at worst exacerbate these existential threats.

Nature **618**, 671-674 (2023)

doi: https://doi.org/10.1038/d41586-023-01979-x

References

1. IPBES. *Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services* (IPBES, 2019).

2. Bateman, I. J. & Balmford, B. *Land Use Policy* **79**, 293–300 (2018).

3. Massimino, D. *et al. BirdTrends 2022: Trends in Numbers, Breeding Success and Survival for UK Breeding Birds* (British Trust for Ornithology, 2023).

4. Lenzen, M. *et al. Nature* **486**, 109–112 (2012).

5. Fuchs, R., Brown, C. & Rounsevell, M. *Nature* **586**, 671–673 (2020).

6. Searchinger, T. D., James, O., Dumas, P., Kastner, T. & Wirsenius, S. *Nature* **612**, 27–30 (2022).

7. Smith, L. G., Kirk, G. J. D., Jones, P. J. & Williams, A. G. *Nature Commun.* **10**, 4641 (2019).

8. Haddad, N. M. *et al. Sci. Adv.* **1**, e1500052 (2015).

9. Balmford, A. *J. Zool.* **315**, 79–109 (2021).

10. Collas, L. *et al. People Nature* **5**, 228–240 (2022).

11. Finch, T. *et al. J. Appl. Ecol.* **58**, 655–666 (2021).

12. The Royal Society. *Multifunctional Landscapes: Informing a Long-Term Vision for Managing the UK's Land* (Royal Society, 2023).

13. Paul, C., Techen, A.-K., Robinson, J. S. & Helming, K. *J. Cleaner Prod.* **227**, 1054–1067 (2019).

14. Furtom, H., Sauer, J. & Jensen, M. S. *Public Choice* **140**, 479–500 (2009).

15. Kahneman, D. *et al. Science* **312**, 1908–1910 (2006).

16. Bateman, I. J. *et al. Science* **341**, 45–50 (2013).

17. Meyfroidt, P., Lambin, E. F., Erb, K.-H. & Hertel, T. W. *Curr. Opin. Environ. Sust.* **5**, 438–444 (2013).

18. Bateman, I. J. *et al. People Nature* **5**, 271–301 (2022).

19. Day, B. & Couldrick, L. *Payment for Ecosystem Services Pilot Project: The Fowey River Improvement Auction* (Univ. East Anglia & Westcountry Rivers Trust, 2013).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01979-x

| Section menu | Main menu |

Work

- **Coming out at work: transgender scientists share their stories** [19 Jun 2023]
  • To mark Pride month, transgender scientists recount transition experiences as political landscapes turn increasingly hostile in many places.
- **How scientists are hacking the genetic code to give proteins new powers** [20 Jun 2023]
  • By modifying the blueprint of life, researchers are endowing proteins with chemistries they've never had before.
- **Teaching robots how to touch** [19 Jun 2023]
  • Through better sensing, Yuhan Hu hopes to bring robots and humans closer together.

- CAREER FEATURE
- 19 June 2023
- Correction 19 June 2023
- Correction 22 June 2023

Coming out at work: transgender scientists share their stories

To mark Pride month, transgender scientists recount transition experiences as political landscapes turn increasingly hostile in many places.

- Sophia Chen

Find a new job



Credit: Vuk Valcic/SOPA Images/LightRocket via Getty

Before coming out as transgender at work, Lisa procured the ultimate academic protection: tenure. A geochemist in her thirties, she found a faculty job at a university in the southwest of the United States. It was 2002, and she knew of many transgender people who had lost their jobs after transitioning. At the time, media depictions cast transgender people largely as freaks, the butt of jokes or hate-crime victims. With job security assured, she told her colleagues she was transitioning — but she did not anticipate their reaction. "The department chair wanted to have a meeting with all the students and announce it," she says. "I was like, 'Do we have to do this?'"

They did. Lisa (not her real name) nodded along as the chair declared her newly public identity to some 50 people — undergrads, graduate students, faculty members and staff. "I guess he thought it was important that everyone know, and [the situation] was not that unusual," says Lisa, who is now 60 and a professor at a university in Massachusetts. "'Some people are trans; some people are French Canadian,' that kind of thing. 'Our colleague is looking a little different. What's going on is completely normal. We all know about this, and we completely support her.'"

Still, the spotlight was a bit much for her. "But it was very nice to see I was being supported," she says. "I didn't expect that."

At the time, stories of transgender scientists, let alone workplaces that were supportive of them, were rare. But if you know where to look, the names of transgender researchers are scattered throughout the scientific record. The British botanist Elke MacKenzie, known for her studies of Antarctic lichen, came out as trans around 1971, at the age of 60. Lynn Conway, a pioneering computer engineer, was fired in 1968 from IBM after she transitioned. (IBM publicly apologized in 2020.) And in the early 2000s, the neuroscientist Ben Barres overheard a colleague praising Barres's work over his sister's, not realizing the 'sister' was Barres before his transition.

"There's this mainstream belief that trans people are constantly brand new, and the issues that affect us are also constantly brand new," says historian Mar Hicks, who researches transgender history at the Illinois Institute of Technology in Chicago. "That's just not true."

But public conversations about gender have become more open — and more scientists are out at work. A 2022 study estimated that at least 1.6 million people in the United States identify as transgender, or as having a gender identity that doesn't match the sex they were assigned at birth (J. L. Herman *et al.* How Many Adults and Youth Identify as Transgender in the United States?; Williams Institute, 2022). Another 2022 study found that 44% of people in the United States know someone personally who is trans. Several countries, including the United States, Canada, Australia, India and Nepal, allow gender options beyond male and female on passports.

Beth Gould, a cosmology PhD student at Newcastle University, UK, works in a research group with three non-binary scientists. "I chose this place, partly because I felt that it was somewhere where being a trans and non-binary physicist was not only tolerated, but actually valued," Gould says.

For Pride month (celebrated in June in the United States and July in the United Kingdom), several transgender scientists shared their academic experiences with *Nature*. Through sharing these stories, they want colleagues to accept their existence and their right to be full members of the scientific community (see 'Transitioning stories').

Transitioning stories: 'The walls go away'

Trans scientists bring perspectives to their work that are as varied as their own backgrounds and coming-out stories.

"Folks have so much more awareness" now, says Shannon Bros, an emeritus ecologist at San Jose State University in California, who transitioned in the late 1990s. "I never even knew what a trans person was growing up."

But trans scientists' stories hold a common theme: to be productive, people must be able to express their true selves at work. As an undergraduate at the University of Minnesota, Minneapolis, in the 1970s, Bros told a therapist about how she would dress as a woman. "This person had me write down what I was doing, on paper, and sign it. Then he told me that what I was doing was against the law," she says. "That's the kind of stuff you had going on then."

After a man sexually assaulted her in her twenties, Bros assumed the persona of a “super macho man” for more than a decade, lifting weights and racing motorcycles. But around the age of 40, she couldn’t put on the act any more. She thought through how to tell the people in her life, but one day in 1997 she threw away the plan and just told a classroom of students that she was a woman. “You live this really bottled-up life, and then all of a sudden, the walls go away, and there’s no way you can put those walls back,” Bros says. After transitioning, she says, she connected better with students.

Being young and trans today is a “varied experience”, says Elijah Beevor, who is pursuing an astrophysics master’s degree at the University of Hertfordshire in Hatfield, UK. Like many young trans people, Beevor has had difficulty in finding gender-affirming health care. In 2018, aged 17, he joined the waiting list to receive hormone therapy through the UK National Health Service and had his first appointment at 22.

Beevor’s family was supportive about his transition, but he has peers whose families don’t accept them. “Some days they’re living as themselves, and some days as their birth gender,” he says.

Jasmine Thomson-Cooke, a physicist at the Dublin Institute for Advanced Studies, identifies as genderless. “It’s just not wanting to be categorized in that way.” Thomson-Cooke describes the label ‘woman’ as an “approximation”, saying: “I don’t know where I fit into that. Part of society welcomes you, as a trans woman, into womanhood, and part of it viciously pushes you out.”

Gabrielle Weber, a physicist at the University of São Paulo in Lorena, Brazil, identifies as a travesti, a gender identity in Latin America that was once a derogatory term, but that the community has reclaimed. She realized aged ten that she was trans when she learned about travestis, who in the past were demeaned and associated with sex work. From a young age, she’d wanted to do science, which seemed incongruous with being a travesti. Now, she says, “calling myself a travesti is me facing my fears”.

Even with more societal awareness, many transgender scientists find workplace support to be patchy and unpredictable. They must navigate

interactions with clumsy and sometimes disrespectful colleagues, and cope with administrative policies that often overlook them and their experiences.

Such obstacles cost a lot of time and energy, says physicist Juani Bermejo-Vega, a postdoc at the University of Granada in Spain. During the COVID-19 pandemic, another researcher tried to kick her out of the women's toilets at work. She's also struggled with the time-consuming process of changing her name on academic papers. An updated name can help to prevent misgendering or misnaming by others, a common harassment technique. It can also help transgender scientists to move forwards with a new identity. "I gave up," she says. "It's more important for me right now during the tenure process that my CV doesn't get messed up."

Trans rights repressed

Despite some progress in certain countries, many transgender scientists work and live in places with an increasingly hostile political landscape regarding trans rights. In the United States, Republican-led state governments are spearheading legislation that targets transgender people. Nine states, including Indiana and Florida, have passed laws making it difficult for schoolchildren to use their preferred pronouns. Florida's and Oklahoma's state governments have made providing gender-affirming care for minors a possible felony. Conservative lawmakers in the United Kingdom are considering redefining the term "sex" in an anti-discrimination law to mean "biological sex". This would make it easier to exclude trans people from certain public spaces and groups, according to the Equality and Human Rights Commission, a public entity advising the government on matters affecting people from sexual and gender minorities.

If you or someone you know is contemplating suicide, or has been sexually assaulted, please reach out. Click for an international directory of helplines and hotlines.

Being transgender is criminalized in at least 13 countries, including the United Arab Emirates, Malaysia and the Gambia, according to a 2020 report. *Nature* did not talk to any transgender scientists from these countries, but the laws can pose complications for transgender scientists who live elsewhere,

too. “More-senior members in my department will suggest that I go to some conference, and I’ll be thinking, ‘It’s illegal to be trans in that country,’ or ‘I do not feel safe as a trans person in this country,’” says Gould. It makes her feel “a bit awkward, but also sad”, she says. “One of the perks of academia is the opportunity to travel to interesting places, and have these amazing experiences, but I feel restricted in my ability to do that.”

Even in more-progressive countries, being trans presents challenges. In Brazil, transgender people can change their legal identity fairly simply — without needing medical documentation, which is required in the United Kingdom, says Gabrielle Weber, a physicist at the University of São Paulo in Lorena, Brazil. However, more trans people have been murdered in Brazil than in any other country for 13 years in a row. The victims are disproportionately Black trans women, as is also the case in the United States.



Gabrielle Weber worked with her university to officially change her identity.Credit: Gabrielle Weber

Still, many trans people choose to transition because it is a matter of life and death. Lisa says that most trans people recognize how difficult transitioning will be before they start the process. “You’ll run into lots of jerks; lots of people aren’t going to support you; you may lose friends; you may lose your job. You had to get to the point where you’re like, ‘I have to do this or I’m going to die.’”

When Shannon Bros, an emerita ecologist at San Jose State University in California, transitioned at around the age of 40, she was married to and had children with a cisgender woman. “With a lot of people, there’s a huge break in trust,” says Bros. “They thought they knew you all your life, and all of a sudden, they didn’t.”

As she rebuilt her work and personal life, Bros found that it was more important to connect emotionally with people than to push facts and “insist on pronouns”. She says that the latter approach was useless. “Most people are never going to understand — that’s a lost cause. You want them to accept you.”

Slipping through the cracks

Even when individual colleagues are supportive, many academic workplaces might overlook trans scientists’ needs because their administrative processes assume two unchangeable genders. For example, universities might not have clear HR protocols for how to handle a trans scientist’s coming out, leading to ad hoc processes similar to Lisa’s debut. Weber had trouble working out how to change her identity officially with the university. An administrator wrote to her, saying: “You are not our first case, but we do not have a standard way of dealing with trans people.” Fortunately, once the university had developed a protocol, her information was updated within a week.

It can be unclear whether gendered scholarships or events welcome transgender students, says Jasmine Thomson-Cooke, a physics PhD student at the Dublin Institute for Advanced Studies. As an undergraduate, Thomson-Cooke wasn’t sure whether it would be acceptable to attend an event for women in physics. The department head e-mailed the organizers

about the issue, but Thomson-Cooke did not go because the clarification took too long.

As the mother of a two-year-old, Bermejo-Vega says that her colleagues do not seem to understand her at-home duties. They often ask her to complete work tasks with short deadlines. “Have people understood that I’m a mum, or do they still see me as a dude and assume that my partner is taking most of the responsibility at home?” she asks.

One challenge, she says, is simply the lack of other trans colleagues. “I have not been able to compare this experience with anyone,” because there are so few trans women — and even fewer trans women who are mothers — in research, says Bermejo-Vega. A 2022 study found that, for transgender undergraduate students who intend to major in science, technology, engineering or mathematics (STEM), the proportion who remain in that track is about 10% less than it is for cisgender students (J. Maloy *et al. CBE Life Sci. Educ.* https://doi.org/kfd3; 2022). “I think many trans people don’t reach the PhD level,” says Bermejo-Vega.



Ecologist Shannon Bros says connecting with colleagues emotionally can help win their acceptance.Credit: Shannon Bros

Both Elijah Beevor, an astrophysics master's student at the University of Hertfordshire in Hatfield, UK, and Thomson-Cooke say that it's important for senior researchers to treat trans students no differently from others. "On the one hand, treat them normally like you would any other student," says Thomson-Cooke. "On the other hand, maybe be more patient, because trans people disproportionately have mental-health problems and might need more time to complete a task." After Brianna Ghey, a 16-year-old transgender girl, was murdered in the United Kingdom in February, Thomson-Cooke and

Gould talked about how “it ruined our productivity for the day”, Thomson-Cooke says.

Creating space for themselves

Transgender scientists have formed communities to socialize and talk about science. A few years ago, Thomson-Cooke started an online Discord community for trans and non-binary physicists, which it runs with Gould. The community organizes mostly virtual social events, such as watching the supernatural television show *Beyond Belief: Fact or Fiction* together. Members hold a recurring event called Thesis Thursdays, in which one of them presents a written introduction to their thesis and fields questions in the chat. “We don’t need to go through the whole rigmarole of, ‘Are you on the pills? Are you getting the surgery?’” says Thomson-Cooke. “We have more interesting things to talk about.”

Bermejo-Vega helped to start a quantum-technology conference, called Q-Turn, that emphasizes diversity and inclusion. Q-Turn held its first event in Florianópolis, Brazil, in 2018, with 114 participants. In November 2020, the conference took place virtually.

Beevor is goalkeeper for TRUK United, a soccer team formed in 2021 that is inclusive of all genders. On 31 March, International Transgender Day of Visibility, Beevor played in the first-ever match to involve a team made up entirely of trans men and trans-masculine players. The opposition was Dulwich Hamlet FC Supporters Team, a cisgender men’s team in south London. His team ultimately lost 8–1, but “considering we literally met each other an hour before, and we were playing against an established team, I think we did pretty well,” says Beevor.

“It’s a gift that we get to experience something that other people don’t,” says Bros. The more acceptance and support that STEM workplaces can show trans colleagues, the more those colleagues will feel safe bringing their whole selves to work. “When I finally did come out, it was like a get-out-of-jail card. I was free to figure out who I was,” says Bros. “People can go their whole lives without doing a lot of introspection. But as a trans person, you have to understand who the heck you are.”

Nature **618**, 871-873 (2023)

doi: https://doi.org/10.1038/d41586-023-01908-y

Updates & Corrections

- **Correction 19 June 2023**: An earlier version of this Career feature misspelt the name of Brianna Ghey.
- **Correction 22 June 2023**: An earlier version of this feature used the wrong pronouns for Jasmine Thomson-Cooke.

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01908-y

- TECHNOLOGY FEATURE
- 20 June 2023

How scientists are hacking the genetic code to give proteins new powers

By modifying the blueprint of life, researchers are endowing proteins with chemistries they've never had before.

- Diana Kwon



A ribosome (centre) builds a new protein (red) from messenger RNA (multicoloured).Credit: Juan Gaertner/Science Photo Library

For most of the history of life on Earth, genetic information has been carried in a code that specifies just 20 amino acids. Amino acids are the building blocks of proteins, which do most of the heavy lifting in the cell; their side-chains govern protein folding, interactions and chemical activities. By limiting the available side chains, nature effectively restricts the kinds of reaction that proteins can perform.

As a doctoral student in the 1980s, Peter Schultz found himself wondering why nature had restricted itself in this way — and set about trying to circumvent this limitation. Several years later, as a professor at the University of California, Berkeley, Schultz and his team managed to do so by tinkering with the machinery of protein synthesis. Although confined to a test tube, the work marked a key early success in efforts to hack the genetic code.

Since then, many researchers have followed in Schultz's footsteps, tweaking the cellular apparatus for building proteins both to alter existing macromolecules and to create polymers from entirely new building blocks. The resulting molecules can be used in research and for the development of therapeutics and materials. But it's been a hard slog, because protein synthesis is a crucial cellular function that cannot easily be changed.

"It's a super-exciting field," says Jason Chin, a synthetic and chemical biologist at the MRC Laboratory of Molecular Biology in Cambridge, UK. "Even from the beginning, I think it was clear that we were doing something really special."

From *vitro* to *vivo*

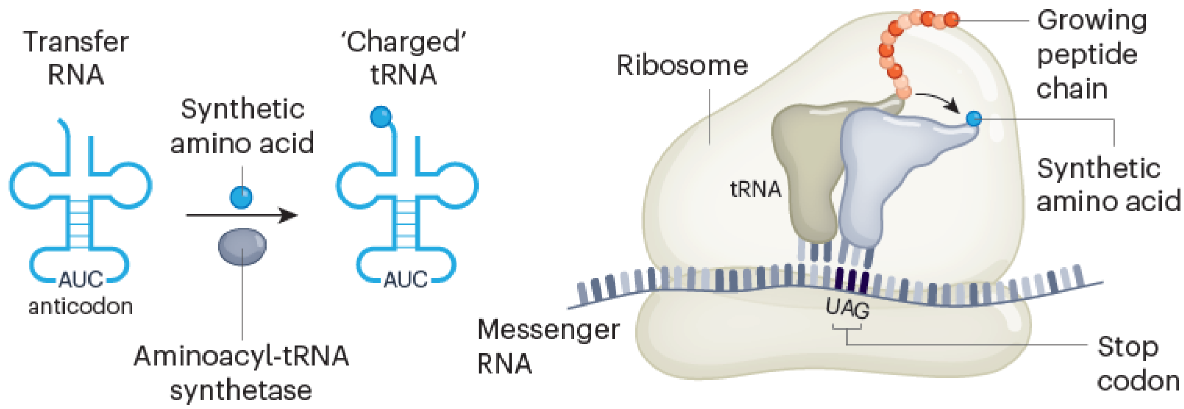
To understand why, we have to consider how proteins are made (see 'Protein hacking').

During transcription, DNA instructions are copied into RNA, which is then translated into proteins in a molecular machine called a ribosome. Each triplet of bases in the RNA message (adenine, cytosine, guanine and uracil)

represents a word, or codon, of the genetic code. There are 64 words in total — 61 that encode for amino acids and 3 that signal the ribosome to stop.

PROTEIN HACKING

By modifying key elements of translation, researchers can give proteins new chemistries. In the most straightforward approach, shown here, a stop codon — a triplet of bases — is repurposed to insert a new amino acid into a peptide chain.



Reprogrammed aminoacyl-tRNA synthetase couples a tRNA to a synthetic amino acid.

The ribosome inserts the synthetic amino acid into a growing peptide chain at a codon that would normally signal translation to terminate.

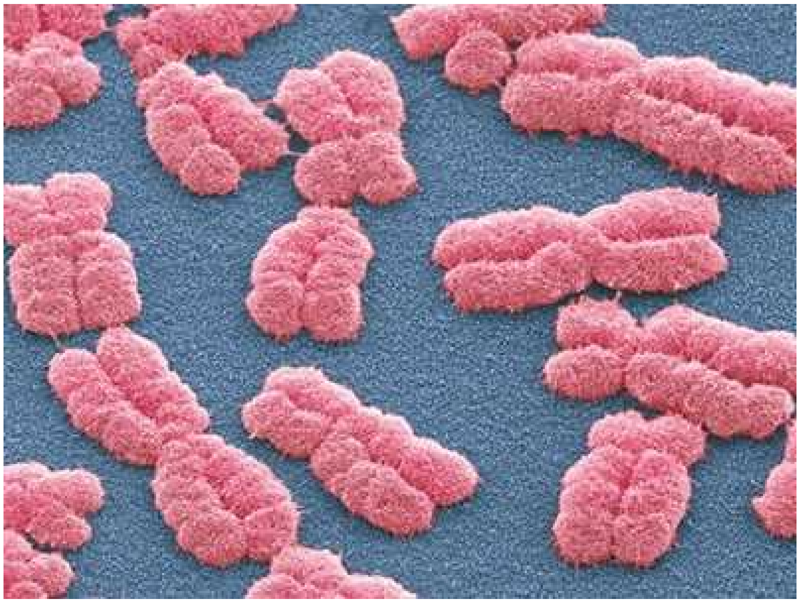
©nature

Amino acids are brought to the ribosome by transfer RNAs (tRNAs), each of which corresponds to a complementary codon. Enzymes known as aminoacyl-tRNA synthetases couple the tRNA to its cognate amino acid.

Making this system accept unnatural, or 'non-canonical', amino acids in a test tube required several key tweaks. First, Schultz's team devised a way to chemically attach unnatural amino acids to a tRNA that could recognize one of the three stop codons. They then introduced a stop codon into the gene for the penicillin-resistance protein, β-lactamase. Using these modifications, the researchers explored how different non-canonical protein variants affected enzyme activity[1].

But that was in a test tube. To incorporate non-canonical amino acids into living cells, the team had to supplement the translational apparatus without altering it. They did so by identifying and modifying tRNAs and aminoacyl-tRNA synthetases that were bio-orthogonal to the host cell — that is, unable

to recognize (and be recognized by) other parts of the cell's translational machinery.



[How to hack the genome]

"People thought that would be very, very difficult to do," says Schultz, who is now the president and chief executive of Scripps Research in La Jolla, California. "Luckily, we were naive enough to think that we could do it." (Schultz has founded several companies, including Ambrx based in La Jolla, which develops biotherapeutics using non-canonical amino acids. He is currently a member of Ambrx's scientific advisory board.)

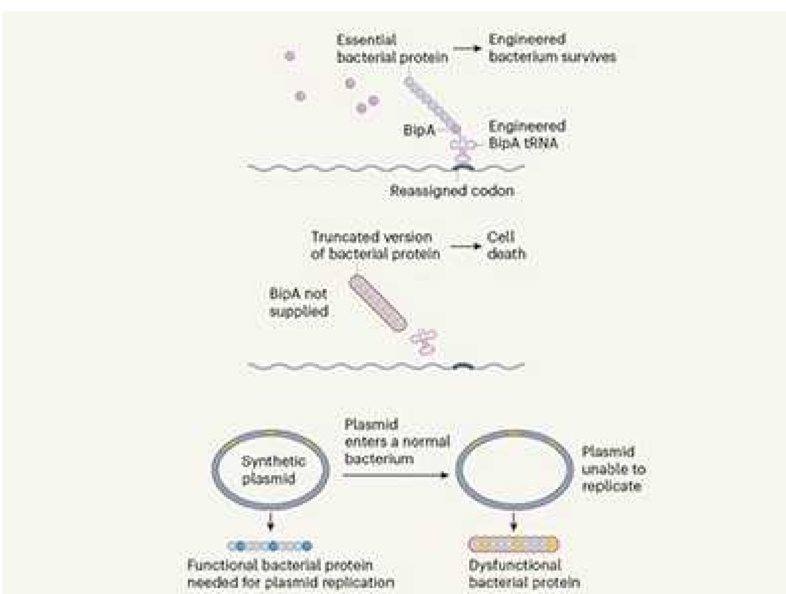
Eventually, Schultz's team settled on a tRNA-synthetase pair from an archaeon, *Methanococcus jannaschii*, which did not readily recognize the tRNAs or the synthetases from the *Escherichia coli* bacteria they planned to modify. The team optimized the molecules to be selective to the new amino acid. In 2001, the researchers engineered this system into *E. coli*, enabling the cells to incorporate non-standard amino acids into proteins using its own translational machinery[2].

"That was probably the beginning of the expansion of the genetic code *in vivo*," says Chin, who was formerly a postdoc in Schultz's laboratory.

A growing toolkit

Schultz's team and others have used this approach to genetically encode more than 200 non-standard amino acids into proteins, providing a powerful tool with which to study protein structure and function. Scientists could, for instance, introduce fluorescent markers or other labels into proteins, or conduct photocaging experiments, in which proteins are rendered inactive ('caged') by chemical groups that can be removed with light.

Scientists have also moved beyond *E. coli* to hack the genetic codes of nematode worms, fruit flies, plants and even mice. Chin, along with Michael Hastings, a neuroscientist at the University of Cambridge, and their colleagues, for instance, hacked the tRNA and tRNA synthetase for the amino acid pyrrolysine to incorporate an amino acid called alkyne lysine N^6-[(2-propynyloxy)carbonyl]-l-lysine (AlkK) into proteins. They used it as a switch to turn genes on and off in mouse brain cells. The researchers deleted a key protein involved in regulating circadian rhythms — daily cycles controlled by the body's internal clock — and introduced a new version of the gene into the cells that could be translated only in the presence of AlkK. Because the mice couldn't make that amino acid on their own, the researchers could control their circadian rhythms by adding or removing the amino acids from the rodents' drinking water[3].

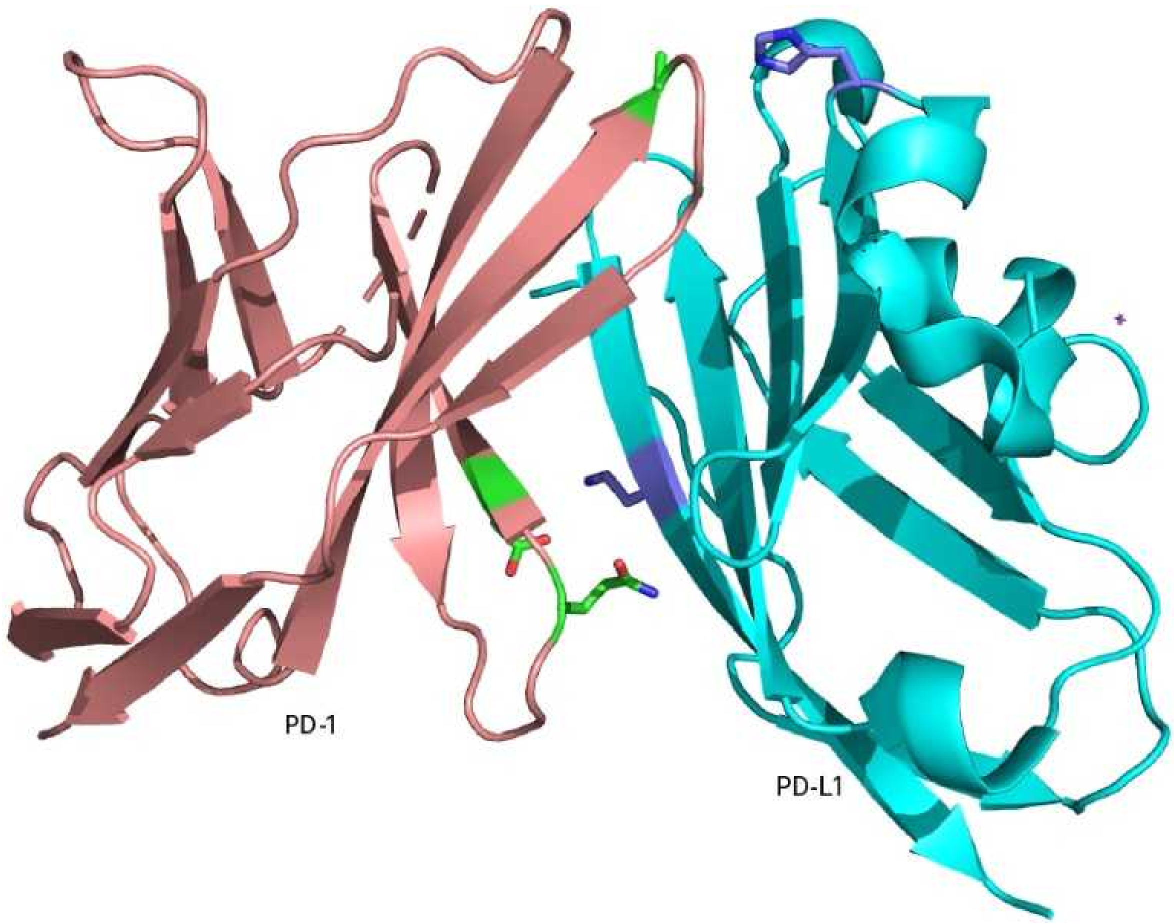


Synthetic bacterial genome upgraded for viral defence and biocontainment

For his part, Schultz has been using the expanded genetic code to ask what organisms would look like if they had more than 20 amino acids to play with. In one study, his team created a library of variants of *E. coli* homoserine O-succinyltransferase, an enzyme involved in the biosynthesis of amino acid methionine that is particularly sensitive to temperature, in which each codon outside the enzyme's catalytic centre was replaced with the non-canonical amino acid, (*p*-benzoylphenyl)alanine (pBzF). The team discovered a variant that was stable at up to 21 °C above the typical range — a feat that they attributed to the formation of a chemical bond between pBzF and another amino acid[4].

Chin, Schultz and other researchers have also begun to apply this technology to the development of therapeutics. Lei Wang, a chemical biologist at the University of California, San Francisco, for instance, is developing protein-based drugs that can form covalent bonds with other biomolecules.

Proteins typically interact with other molecules through relatively weak, non-covalent interactions, but covalent bonding could enhance their potency, says Wang. In 2020, his team incorporated the non-canonical amino acid fluorosulfate-l-tyrosine (FSY) into PD-1, an immune checkpoint protein that helps to rein in the body's immune response, to create an anti-tumour drug. Typically, the interaction between PD-1 on T cells, and PD-L1 on tumour cells, dampens the immune response, allowing the tumour to escape immune surveillance. When Wang's team injected the FSY-containing PD-1 into mice engrafted with human cancer cells, the protein formed an irreversible covalent bond with PD-L1, causing the tumours to shrink[5]. Wang is a scientific adviser at the company Enlaza Therapeutics in La Jolla, which has developed this therapeutic strategy.



PD-1 (left), an immune checkpoint protein, bound to its ligand PD-L1.Credit: Q. Li *et al. Cell* **182**, 85–97 (2020)

Meanwhile, in Beijing, Tao Liu, a chemical and synthetic biologist at Peking University, and his team have been applying genetic expansion to cell and gene therapies. In a 2021 study, Liu and his team reported engineering cells that, in the presence of the synthetic amino acid *O*-methyl tyrosine, express insulin[6]. When implanted into diabetic mice, the researchers could control the animal's blood glucose levels by controlling how much *O*-methyl tyrosine they dispensed in the animals' food.

Expanding applications

Beyond encoding new amino acids, one application of an extended genetic code is genetic isolation. With 61 codons for 20 amino acids, the genetic code is redundant, meaning that multiple codons encode the same amino

acid. It is also nearly universal. By replacing all instances of a given codon for a synonymous one and removing the machinery that uses the original codon, researchers can render the cell effectively immune to foreign DNA — including pathogens.

In a 2022 study Chin and his group demonstrated this concept; they created an *E. coli* mutant in which two of its six serine codons were reassigned to encode for other amino acids — then they deleted the tRNAs that recognized the original serine codons. They further tweaked the system to ensure that viruses couldn't use their own tRNAs, if they had them. The resulting cells were resistant to horizontal gene transfer from other bacteria, as well as to viral infection[7]. Another group, led by synthetic biologist George Church at Harvard University in Cambridge, Massachusetts, applied a similar 'genetic firewall' approach to create phage-resistant bacteria in a paper published earlier this year[8].

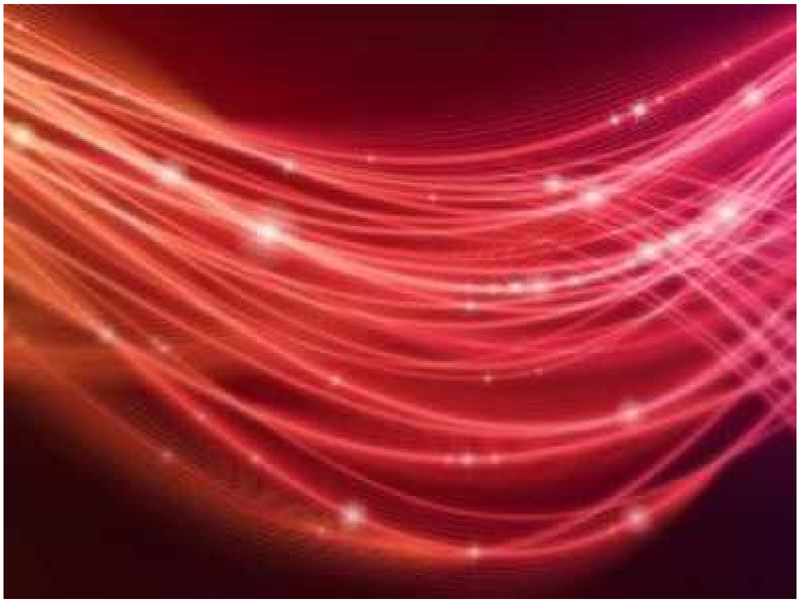
Researchers can also use a modified genetic code to create polymers. In 2021, Chin's team hacked the genetic code to synthesize short polymers, and even an artificial circular structure called a macrocycle, in *E. coli*[9]. Now, Chin hopes to push this technology further to create cell factories that can churn out entirely new polymers, such as plastics. Like proteins, plastics are composed of long strings of monomers. But, although the genetic code dictates the sequence of amino acids in proteins, no equivalent system exists for artificial polymers.

"If you could encode the expanded set of building blocks in the same way that we can proteins, then we could turn cells into living factories for the encoded synthesis of polymers for everything from new drugs to materials," says Chin. (Chin founded a company, Constructive Bio based in Saffron Walden, UK, to advance this goal.)

Pushing boundaries

Besides codon reassignment, researchers can also increase the number of available protein building blocks by expanding the nucleic acid alphabet from four bases to six, thus increasing the number of possible triplet codons to 216. In 2014, a team led by biochemist Floyd Romesberg, who was then

at Scripps Research, reported creating a bacterial strain with a six-base genetic alphabet that could successfully replicate[10]. The group subsequently demonstrated that these cells could use their expanded DNA to produce proteins containing non-canonical amino acids[11].



NatureTech hub

Another approach is to expand the length of a codon from three bases to four, thus increasing the number of possible codons to 256. This requires modifying multiple pieces of the translation machinery, including the ribosome. Chin and his team have used this strategy to incorporate four non-canonical amino acids into *E. coli* on a transcript that also contains conventional three-base codons[12]. Other researchers have been exploring whether it might even be possible to create an all-quadruplet genetic code.

Some researchers are attempting more extreme alterations. These include backbone modifications — creating so-called β- or γ-amino-acids (as opposed to α-amino-acids, which are found in nature), or amino acids that are reverse mirror images of standard amino acids. Polymers built from either type of building block would probably be highly stable, because typical protein-degrading machinery would not be able to recognize them.

But existing translational machinery isn't built to accept these exotic amino acids, including the aminoacyl-tRNA synthetases that attach amino acids

onto tRNA. Hiroaki Suga, a chemical biologist at the University of Tokyo, developed one workaround in 2006 with the ‘flexizyme’, an RNA-based catalyst that can perform the job of protein synthetases: linking amino acids to tRNAs[13].

Suga focuses on synthesizing polymers *in vitro*, because it provides the greatest flexibility in terms of the modifications he can make. Using the flexizyme, his team combined up to 11 non-canonical amino acids with 12 standard amino acids in a single macrocycle[14]. Suga co-founded the biotechnology company PeptiDream, based in Kawasaki, Japan, to develop drugs using the same technology.

Again, however, that work is *in vitro*; applying backbone- or configuration-modified amino acids to proteins in cells remain a challenge. Although there are a few instances in which scientists have been able to incorporate these more exotic amino acids, further steps are necessary to make the process more efficient. As well as finding synthetases that could work with these amino acids *in vivo*, in many cases scientists would have to design a ribosome that can process these new amino acids — while still carrying out its usual duties.

In two recent studies Alanna Schepartz, a chemical and synthetic biologist at the University of California, Berkeley, and her team report steps toward solving these problems. In one of the papers, they describe a synthetase from the archaeon *Methanomethylophilus alvus* that can accept non-α-amino-acids and is bio-orthogonal to *E. coli.* In the other study, the team reports a computational technique for screening unique backbone monomers that the *E. coli* ribosome can process efficiently. The researchers say this will help to identify non-α-amino-acids that are most likely to be compatible with an existing ribosome[15,16].

Other groups are working to develop a ribosome that can produce proteins with exotic amino acids from scratch. Last October, researchers at Westlake University in Hangzhou, China, reported a mirror-image RNA polymerase that could synthesize all the RNA molecules that are needed to produce a mirror-image ribosome[17]. Although many more steps are needed to make a mirror-image ribosome a reality, creating such a molecule would be an

important step towards using the translational machinery to make mirror-image proteins.

Other researchers are working to reprogram the ribosome to create polymers with carbon–carbon bonds, as opposed to the nitrogen–carbon amide bonds they typically forge to link amino acids.

Should these strategies come to fruition, they would endow scientists with tremendous synthetic power over new polymers. Nobody knows what properties can emerge from a synthetic polymer that's created with the same length and level of sequence definition as a protein, "because no such molecules have ever been made", Schepartz says. But she and others hope to get there soon. "It's a very exciting time for the field."

Nature **618**, 874-876 (2023)

doi: https://doi.org/10.1038/d41586-023-01980-4

References

1. Noren, C. J. Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. *Science* **244**, 182–188 (1989).
2. Wang, L. Brock, A., Herberich, B. & Schultz, P. G. *Science* **292**, 498–500 (2001).
3. Maywood, E. S. *et al. Proc. Natl Acad. Sci. USA* **115**, E12388–E12397 (2018).
4. Li, J. C., Liu, T., Wang, Y., Mehta, A. P. & Schultz, P. G. *J. Am. Chem. Soc.* **140**, 15997–16000 (2018).
5. Li, Q. *et al. Cell* **182**, 85–97 (2020).
6. Chen, C. *et al. Nature Chem. Biol.* **18**, 47–55 (2022).
7. Zürcher, J. F. *et al. Science* **378**, 516–523 (2022).

8. Nyerges, A. *et al. Nature* **615**, 720–727 (2023).

9. Robertson, W. E. *et al. Science* **372**, 1057–1062 (2021).

10. Malyshev, D. A. *et al. Nature* **509**, 385–388 (2014).

11. Zhang, Y. *et al. Nature* **551**, 644–647 (2017).

12. Dunkelmann, D. L., Oehm, S. B., Beattie, A. T. & Chin, J. W. *Nature Chem.* **13**, 1110–1117 (2021).

13. Murakami, H., Ohta, A., Ashigai, H. & Suga, H. *Nature Methods* **3**, 357–359 (2006).

14. Peacock, H. & Suga, H. *Trends Pharma. Sci.* **42**, 385–397 (2021).

15. Fricke, R. *et al. Nature Chem*. https://doi.org/10.1038/s41557-023-01224-y (2023).

16. Watson, Z. L. *et al. Nature Chem*. https://doi.org/10.1038/s41557-023-01226-w (2023).

17. Xu, Y. & Zhu, T. F. *Science* **378**, 405–412 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01980-4

- WHERE I WORK
- 19 June 2023

Teaching robots how to touch

Through better sensing, Yuhan Hu hopes to bring robots and humans closer together.

- James Mitchell Crow



Yuhan Hu is a PhD student researching mechanical engineering at Cornell University, New York. Credit: Jesse Winter for *Nature*

Touch is the first sense that humans develop before birth. It is an intimate, emotional way to communicate and can convey a lot of information.

I study touch in the context of human–robot interaction, for my PhD programme at Cornell University in Ithaca, New York. Communication through touch is important for social and companion robots.

We developed a soft robot 'skin' that enables touch-based interaction. Our robots can communicate through alterations to the the shape, size and motion of textures on their skin. We can create goosebumps, like those that appear on the skin of someone who's excited. We can also create spikes, inspired by porcupinefish, which puff into a spiky ball when they're angry.

In this photograph, I am holding a robot skin-wrinkle module, which mimics the way a human forehead can wrinkle up. We have made tentacled robots, like the blue robot in the picture. It brings together sighted and visually impaired children for a storytelling activity; each child touches the tentacles on a robot arm, and the tentacles' movements are mapped to the emotions of a character.

We are exploring a texture-changing steering wheel for autonomous cars, to enable a more emotional connection between driver and car. Driving is usually a rigid human–machine interaction. By contrast, when you are riding a horse, you can feel its emotions through the physical connection, which we're trying to recreate: the steering wheel might convey calm on a quiet road, and focus in heavy traffic.

As companion robots, such as the striped robot in the photograph, move into our homes, their camera vision systems might raise privacy concerns. We have developed a mobile robot that maintains privacy by relying on touch. The robot's camera is inside its body, beneath a translucent skin that clouds its view. When a hand is placed on the robot's skin, the camera can recognize gestures, from a 'go away' tap to a 'move closer' stroke, from the hand's shadow.

Nature **618**, 878 (2023)

doi: https://doi.org/10.1038/d41586-023-01981-3

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01981-3

Research

- **A metal-poor star with abundances from a pair-instability supernova** [07 Jun 2023]
  • The chemical composition of the Galactic halo star J1010+2358 shows extremely low sodium and cobalt abundances, different from most other halo stars, indicative of a very metal-poor star being seeded with elements from a pair-instability supernova.
- **Density-wave ordering in a unitary Fermi gas with photon-mediated interactions** [24 May 2023]
  • An experiment with atoms and photons provides a fully tunable and microscopically controllable platform for the experimental study of the interplay of superfluidity and density-wave order.
- **High-brightness scalable continuous-wave single-mode photonic-crystal laser** [14 Jun 2023]
  • We developed large-scale photonic-crystal surface-emitting lasers with controlled Hermitian and non-Hermitian couplings inside the photonic crystal and a pre-installed spatial distribution of the lattice constant, which leads to the realization of a continuous-wave brightness of 1 GW cm−2 sr−1.
- **Entropic repulsion of cholesterol-containing layers counteracts bioadhesion** [21 Jun 2023]
  • Entropic repulsion caused by interfacial orientational fluctuations of cholesterol layers restricts protein adsorption and bacterial adhesion, providing a conceptually new physicochemical perspective on biointerfaces that may guide future material design in regulation of adhesion.
- **Increasingly negative tropical water–interannual CO2 growth rate coupling** [31 May 2023]
  • Records show that tropical water availability rather than temperature appears to have been increasingly controlling the interannual variability of the terrestrial carbon cycle over the past 59 years.
- **Polygenic scoring accuracy varies across the genetic ancestry continuum** [17 May 2023]
  • Using two large biobank datasets, a study shows that the accuracy of polygenic scores decreases as a function of relatedness at the individual level when modelling genetic ancestry as a continuum.

- **The illusion of moral decline** [07 Jun 2023]
  • We show that the perception of moral decline is pervasive, perdurable, unfounded and easily produced, and suggest that this illusion has implications for research on the misallocation of scarce resources, the underuse of social support and social influence.
- **Psychedelics reopen the social reward learning critical period** [14 Jun 2023]
  • Behavioural electrophysiological and transcriptomic studies in mice show that psychedelic drugs reopen the social reward learning critical period and suggest that this involves reorganization of the extracellular matrix.
- **A tripartite rheostat controls self-regulated host plant resistance to insects** [14 Jun 2023]
  • Insect salivary protein (BISP) targets OsRLCK185 to suppress defence in susceptible plants, whereas in resistant plants BISP binds BPH14 to activate host plant resistance. To restore cellular homeostasis, the resistance mechanism is fine-tuned by selective autophagy.
- **Signalling by senescent melanocytes hyperactivates hair growth** [21 Jun 2023]
  • Senescent melanocytes of skin nevi drive hyperactivation of hair growth through the signalling factor SPP1.
- **In situ tumour arrays reveal early environmental control of cancer immunity** [31 May 2023]
  • Skin tumour array by microporation (STAMP) captures the dynamic relationships of spatial, cellular and molecular components of tumour rejection and has the potential to translate therapeutic concepts into successful clinical strategies.
- **Ultraviolet radiation shapes dendritic cell leukaemia transformation in the skin** [07 Jun 2023]
  • Blastic plasmacytoid dendritic cell neoplasm (BPDCN) arises from clonal (premalignant) haematopoietic precursors in the bone marrow, and BPDCN skin tumours first develop at sun-exposed anatomical sites and are distinguished by clonally expanded mutations induced by ultraviolet radiation.
- **Engineered tRNAs suppress nonsense mutations in cells and in vivo** [31 May 2023]
  • Suppressor tRNAs adapted to the amino acid that they carry enable readthrough of premature termination codons introduced by nonsense mutations and show potential for the treatment of genetic diseases such as cystic fibrosis.
- **A cytosolic surveillance mechanism activates the mitochondrial UPR** [07 Jun 2023]

• We identify a highly controlled cytosolic surveillance mechanism that integrates independent mitochondrial stress signals to initiate the mitochondrial unfolded protein response (UPR), revealing a link between mitochondrial and cytosolic proteostasis.

- **Genome expansion by a CRISPR trimmer-integrase** [14 Jun 2023]
  • CRISPR systems lacking Cas4 can use fused or recruited exonucleases for faithful acquisition of new CRISPR immune sequences.

- **Structural basis for FGF hormone signalling** [07 Jun 2023]
  • This study reveals how Klotho and heparan sulfate glycosaminoglycan coreceptors enable FGF hormones to induce asymmetric 1:2 FGF–FGFR dimerization mandatory for FGFR kinase activation and hence signalling.

- Article
- Open Access
- Published: 07 June 2023

A metal-poor star with abundances from a pair-instability supernova

- Qian-Fan Xing,
- Gang Zhao ,
- Zheng-Wei Liu,
- Alexander Heger,
- Zhan-Wen Han,
- Wako Aoki,
- Yu-Qin Chen,
- Miho N. Ishigaki,
- Hai-Ning Li &
- …
- Jing-Kun Zhao

Show authors

Nature **volume 618**, pages 712–715 (2023)

- 11k Accesses
- 347 Altmetric
- Metrics details

Abstract

The most massive and shortest-lived stars dominate the chemical evolution of the pre-galactic era. On the basis of numerical simulations, it has long

been speculated that the mass of such first-generation stars was up to several hundred solar masses[1,2,3,4]. The very massive first-generation stars with a mass range from 140 to 260 solar masses are predicted to enrich the early interstellar medium through pair-instability supernovae (PISNe)[5]. Decades of observational efforts, however, have not been able to uniquely identify the imprints of such very massive stars on the most metal-poor stars in the Milky Way[6,7]. Here we report the chemical composition of a very metal-poor (VMP) star with extremely low sodium and cobalt abundances. The sodium with respect to iron in this star is more than two orders of magnitude lower than that of the Sun. This star exhibits very large abundance variance between the odd- and even-charge-number elements, such as sodium/magnesium and cobalt/nickel. Such peculiar odd–even effect, along with deficiencies of sodium and α elements, are consistent with the prediction of primordial pair-instability supernova (PISN) from stars more massive than 140 solar masses. This provides a clear chemical signature indicating the existence of very massive stars in the early universe.

Main

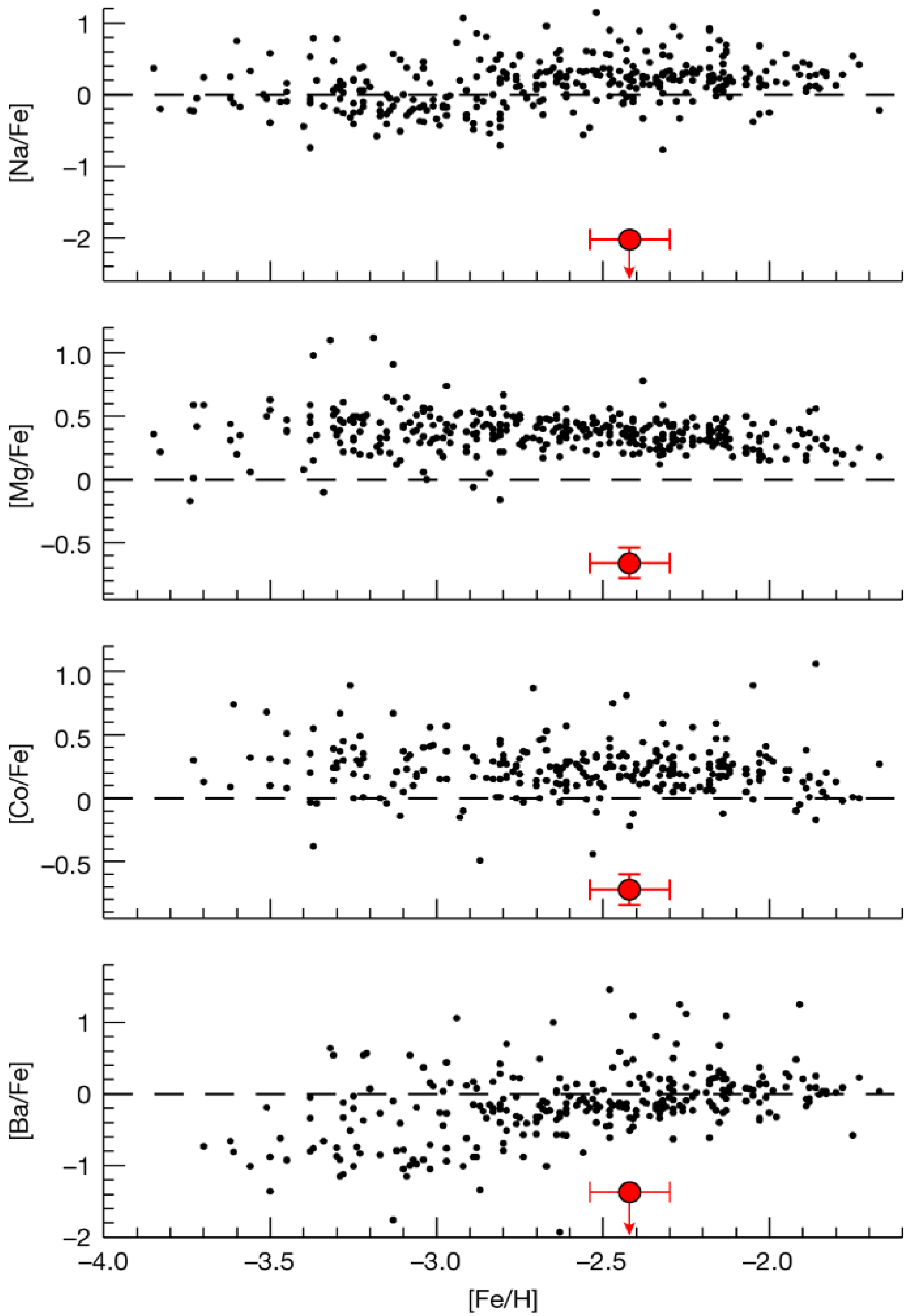
The Galactic halo star LAMOST J1010+2358 (hereafter, J1010+2358, V-band magnitude $V = 16.01$) was identified as a VMP star to have a relatively low Mg abundance based on the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) survey[8,9]. The analysis of the high-resolution spectrum from follow-up observation with the Subaru Telescope (Methods) confirms that J1010+2358 is a VMP star ([Fe/H] = −2.42) with extremely low α-element abundances (for example, [Mg/Fe] = −0.66). More than 400 VMP stars have been identified from the LAMOST survey and follow-up observations with high-resolution spectra[10,11]. None of these VMP stars exhibits such low α-element abundances. The remarkably low α elements to iron ratios, along with the unusual absence of sodium and barium, indicate that J1010+2358 may have recorded a chemical enrichment history completely different from those of most halo stars.

The abundances of Mg, Si, Ca, Ti, Cr, Mn, Fe, Co and Ni shown in Table 1 are determined from the equivalent widths (EWs) based on one-dimensional plane-parallel local thermodynamic equilibrium (LTE) model atmospheres[12]. The upper limits of Na, Sc, Zn, Sr and Ba abundances are estimated by the

spectrum synthesis method. As a VMP star with [Fe/H] = −2.42, the chemical abundances of J1010+2358 are very peculiar in comparison with other metal-poor stars in the Milky Way. This star has sub-solar [X/Fe] ratios for Na, Mg, Ca, Ti, Cr, Mn, Co, Ni and Zn. Its Na to Fe ratio ([Na/Fe] < −2.02) is lower than 1/100th of the solar value [29], whereas almost all other metal-poor stars exhibit Na/Fe ratios ([Na/Fe] > −1) greater than 1/10th of the solar value (Fig. 1). Furthermore, the Mg to Fe ratio of J1010+2358 ([Mg/Fe] = −0.66) is substantially lower than the typical abundance ratio of the Galactic halo stars with similar metallicities. The abundance of Co in this star is unusually low for its metallicity. What stands out is the large variance between the odd-*Z* and even-*Z* elemental abundances, the so-called odd–even effect, such as Na/Mg and Co/Ni. The absence of absorption lines of neutron-capture elements such as Sr and Ba in J1010+2358 is also notable. The upper limits of the abundances of Sr and Ba are lower than expected for a VMP star. This implies that there is no evidence for enrichment of rapid or slow neutron-capture process elements[13].

Table 1 Abundance results obtained for J1010+2358

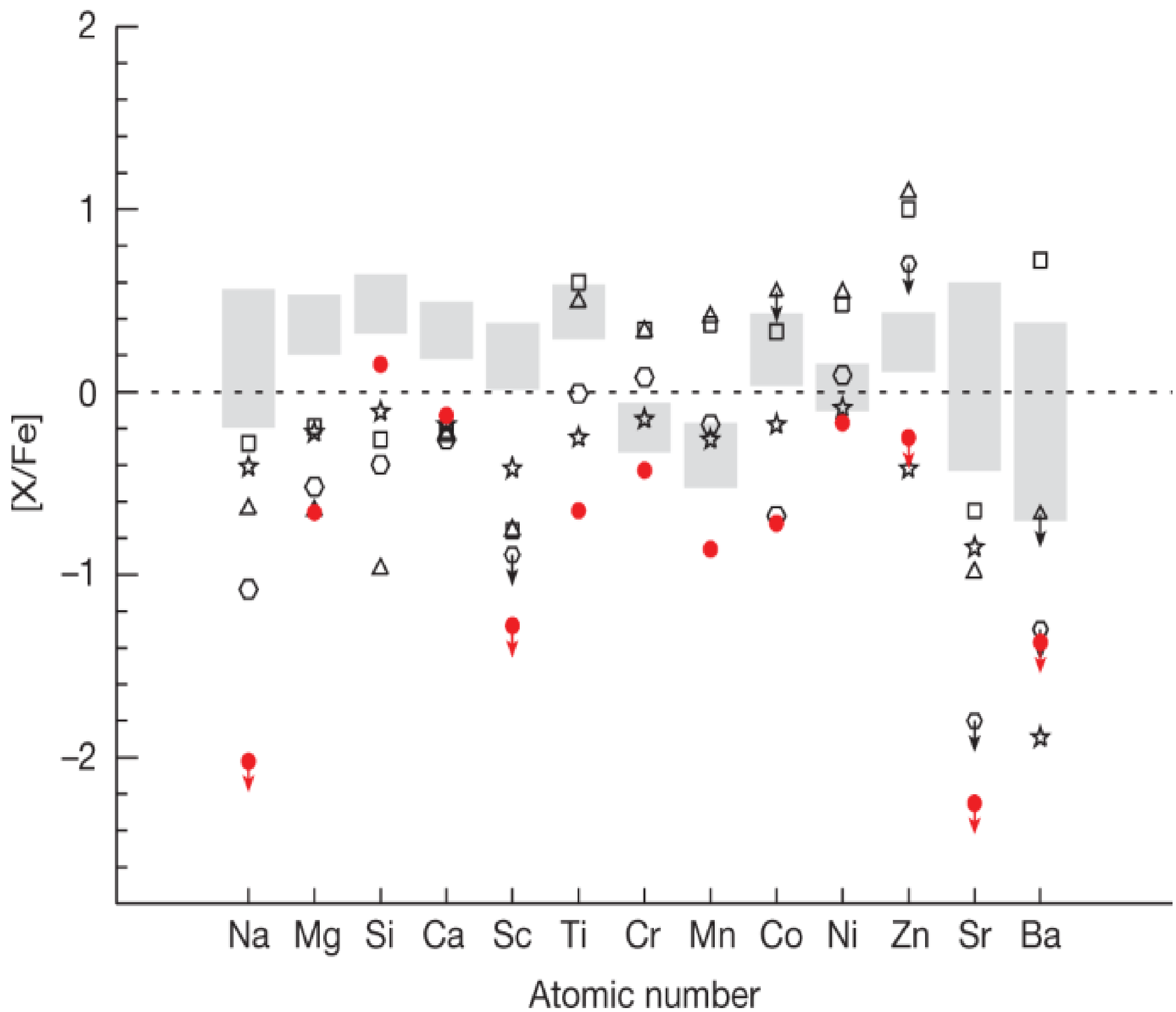
Fig. 1: The abundances of J1010+2358 in comparison with those of other metal-poor stars.


[Na/Fe]
[Mg/Fe]
[Co/Fe]
[Ba/Fe]
[Fe/H]
−4.0
−3.5
−3.0
−2.5
−2.0

J1010+2358 is shown as the red circles. The black circles indicate the metal-poor stars from the literature[10,11]. The arrows indicate the upper limits. The error bars are 1σ uncertainties of the observed abundances.

The metal-poor stars in the Galactic halo typically possess enhanced α-element abundances ([Mg/Fe] > +0.3) owing to the chemical enrichment with core-collapse supernovae (CCSNe, enhancement of α elements) and the absence of Type Ia supernova (SN Ia) contributions[14] (enhancement of iron). The low abundances of α elements with respect to iron in J1010+2358 show an excessive enrichment of iron. A few metal-poor stars are known to have low α element to iron ratios (α-poor stars)[15,16] that are similar to J1010+2358, but none of these stars exhibits such low abundances of iron peak [X/Fe] (for example, Cr, Mn, Co, Ni and Zn) as J1010+2358 (Fig. 2). The model at present [14,16,17] is that the abundance patterns of previously known α-poor stars are the result of large iron yields from SN Ia. Combined with the enrichment of α elements (for example, Mg, Si and Ca) by CCSNe[18], the contribution of SN Ia leads to the increase of iron-peak elements only and, thereby, to the decrease of [α/Fe] ratio[19]. As shown in Fig. 2, the previously known α-poor stars present normal or higher abundances of [Cr/Fe] and [Mn/Fe], along with low α element to iron ratios. By contrast, the abundances of [Cr/Fe] and [Mn/Fe] in J1010+2358 are much lower than those of other stars, ruling out any contribution from SN Ia. In general, the peculiar abundance pattern of J1010+2358 is markedly different from any known stars. Its abundance pattern is not likely to be produced by nucleosynthetic yields of several progenitors, as contributions from normal nucleosynthesis (for example, core-collapse supernova (CCSN) or SN Ia) would obscure such a peculiar feature of chemical abundances. The entire abundance pattern could be produced by nucleosynthesis from a very massive first-generation star, which contributes excess iron into the interstellar medium by means of a PISN[5].

Fig. 2: Abundance pattern of J1010+2358.

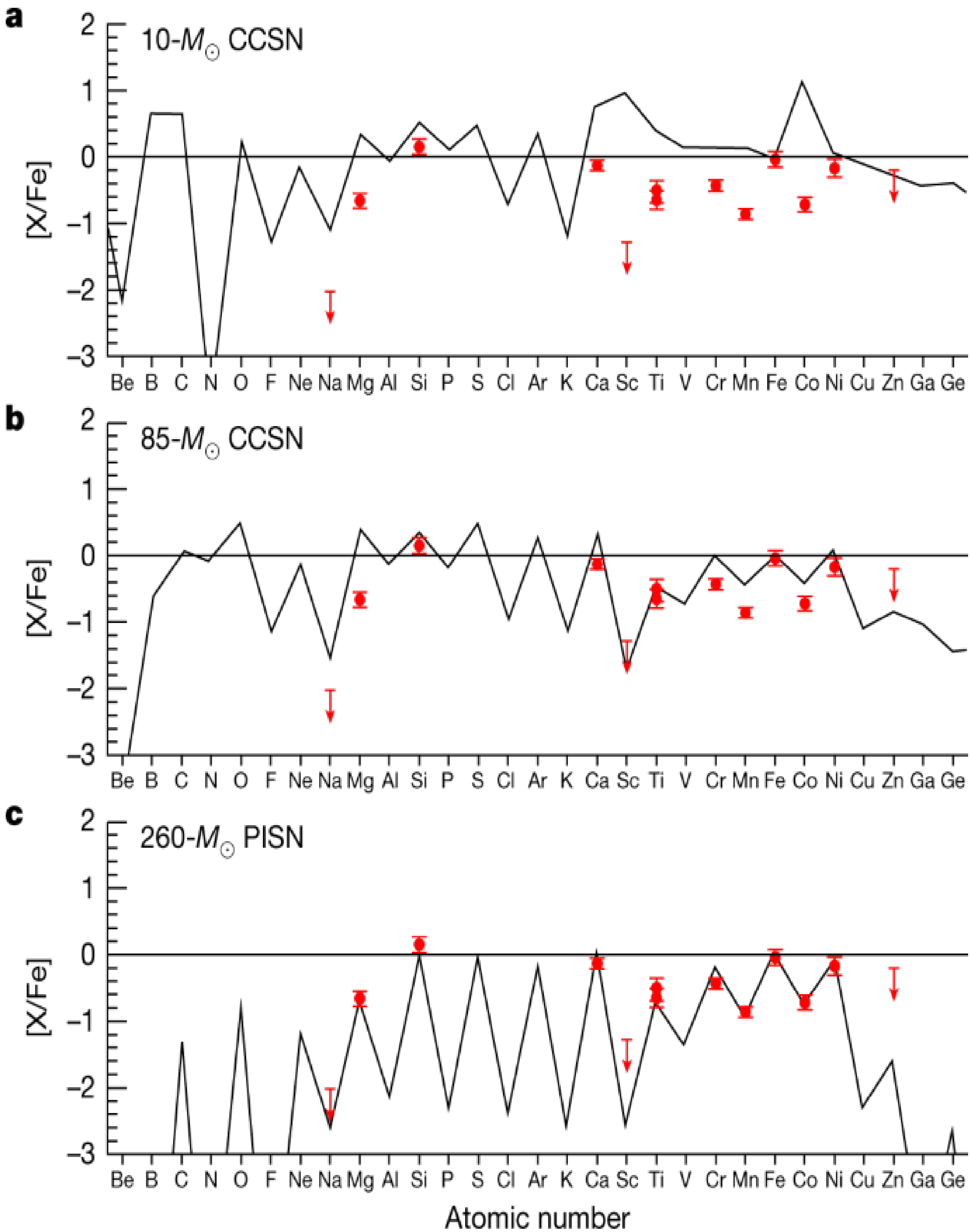


The red circles denote J1010+2358. The open symbols indicate four previously known metal-poor stars with sub-solar [Mg/Fe] ratios. The abundances of these α-poor metal-poor stars ($-2.46 \leq$ [Fe/H] ≤ -1.91) have been well studied on the basis of high-resolution spectroscopic analysis[15,16]. The shaded regions indicate abundances of other metal-poor stars from the literature[10,11]. The arrows represent the upper limits.

We compare the observed abundance pattern of J1010+2358 with theoretical predictions on nucleosynthesis yields of CCSNe and PISNe (Fig. 3). The evolution of massive first stars with initial masses of about 10–140 $M_{\odot}$ is considered to lead to iron-core collapse at the end to explode as CCSNe. The non-rotating stars with helium core masses of about 65–130 $M_{\odot}$ (corresponding to initial masses of zero-age main-sequence massive stars of 140–260 $M_{\odot}$) are expected to lead to the production of electron–positron

pairs (e^+/e^-) before oxygen ignition, causing rapid contraction and the ignition of explosive oxygen burning. This process finally leads to an energetic thermonuclear runaway that is referred as a PISN, ejecting a large amount of heavy elements and leaving no remnant behind. The PISNe require their progenitors to have a helium core with a mass greater than 65 $M_\odot$, which can only be fulfilled at extremely low metallicity because a very massive star triggers too strong stellar wind mass loss at high metallicity to form a massive helium core.

Fig. 3: Comparison of observed abundances and models.



The chemical abundances of J1010+2358 compared with the predictions from three theoretical supernova models[5,18]: a 10-$M_\odot$ CCSN (**a**); a 85-$M_\odot$

CCSN (**b**); a 260-$M_\odot$ PISN with a 130-$M_\odot$ He core (**c**). The error bars are 1σ uncertainties of the observed abundances.

As shown in Fig. 3, the observed ratio of [Mg/Fe] and the upper limit of [Na/Fe] in J1010+2358 are much smaller than those from theoretical predictions of CCSNe, which exclude the possibility that the observed abundance pattern of J1010+2358 results from nucleosynthesis of a CCSN. Furthermore, the low [Cr/Fe], [Mn/Fe] and [Co/Fe] ratios are inconsistent with the expectation of CCSNe. The PISN models[5] for zero-metallicity non-rotating progenitors with initial masses of 140–260 $M_\odot$ that were computed by using the one-dimensional implicit hydrodynamics code KEPLER[20,21] are also compared with the abundance pattern of J1010+2358 for exploring the enrichment source. The nucleosynthetic yields predicted by the PISN model with an initial mass of 260 $M_\odot$ can closely reproduce the observed abundance pattern of J1010+2358 (Fig. 3). The odd–even effect in PISNe is substantially larger than that in CCSNe (ref. [22]), which agrees well with the fact that J1010+2358 exhibits a stronger odd–even effect compared with other metal-poor stars in the Milky Way. The absence of neutron-capture elements in J1010+2358 is also in good agreement with the expectation of low-metallicity PISNe. Given the lack of heavy seed nuclei and appreciable neutron sources in helium burning, no *s*-process is expected in low-metallicity PISNe. Also, the production of *r*-process elements requires a very neutron-rich condition. The neutron deficiency is predicted in PISNe, which leads to the lack of *r*-process in PISNe. The low abundance ratios of [Na/Fe], [Mg/Fe], [Mn/Fe] and [Co/Fe] seen in J1010+2358 strongly suggest a PISN contribution. The iron-peak elements in PISNe are mostly produced by the incomplete Si burning and the incomplete-Si-burning regions in PISNe are much smaller than those in CCSNe, which leads to a low production of Mn and Co in PISNe. The production of Na requires excess neutrons and it is very sensitive to the initial metallicity in PISNe. The neutronization during the final evolutionary stages in PISNe is much less notable than in CCSNe, leading to a remarkable deficiency of odd-charged nuclei compared with even-charged nuclei in the nucleosynthesis of PISNe. The yield of α elements such as Mg is expected to be inefficient for massive, low-metallicity PISN progenitor models.

The discovery of J1010+2358 has provided a clear chemical signature for the existence of PISNe from very massive first stars. Its metallicity ([Fe/H] = −2.42) shows that the second-generation stars formed in the material enriched by supernovae from the first massive metal-free stars do not have to be extremely metal poor ([Fe/H] < −3)[23,24,25]. The extremely metal-poor stars are formed in the pristine gas polluted by very few CCSNe from Population III stars[7,26] with masses less than 100 $M_{\odot}$. Given that such Population III stars (<100 $M_{\odot}$) live longer than the progenitors (140–260 $M_{\odot}$) of PISNe, the second-generation stars with relatively high metallicities ([Fe/H] > −3), such as J1010+2358, should be formed in PISN-dominated cloud before the birth of the most metal-poor stars with CCSN imprints. Notably, a very low [Mg/Fe] as found for J1010+2358 has been observed in a broad line region in a very-high-redshift quasar[27] with a high [Fe/H], for which a large amount of iron contributed by PISNe is suggested. The peculiar abundances of J1010+2358 provide key features for identifying PISN signatures. Detailed studies of VMP stars included in the large stellar abundance databases[28] will facilitate the discovery of more PISN-dominated stars and provide an essential clue to constraining the initial mass function in the early universe.

Methods

Observational data and stellar parameters

The high-resolution (R = 36,000) spectroscopic observation of LAMOST J1010+2358 was obtained by using Subaru/High Dispersion Spectrograph (HDS)[30] on 3 June 2015. The high-resolution spectra cover the wavelength range of 4,000–6,800 Å with a gap of 5,330–5,430 Å. The resolving power of $R \approx 36{,}000$ is obtained by using a 1.0-arcsec slit and 2 × 2 CCD pixel binning. The signal-to-noise ratios at 4,300 and 5,000 Å are 50 and 70, respectively. Data reduction, including bias correction, flat fielding, scattered light subtraction and wavelength calibration, was carried out with the IRAF echelle package.

The radial velocity of J1010+2358 was measured from Fe I lines that are used for abundance analysis. The heliocentric radial velocity derived from

the high-resolution spectra, −101.8 ± 0.7 km s^{-1}, are in good agreement with that from the LAMOST spectra. The EWs of isolated absorption lines were measured by fitting Gaussian profiles with the IRAF task splot by using a line list compiled from the literature[31,32]. The kinematic analysis indicates that this star is a Galactic halo star on a retrograde orbit.

Stellar parameters, including effective temperature (T_{eff}), surface gravity (log g) and microturbulent velocity (v_t), are determined spectroscopically from isolated absorption lines of Fe based on the LTE model atmospheres[12]. The abundances of individual Fe I and Fe II lines are derived by using the MOOG program[33]. T_{eff} is determined by forcing the abundances derived from individual Fe I lines to be independent of their excitation potential. We also estimate T_{eff} from the $(V–K)_0$ colours[34] (T_{eff} = 5,810 K), which is in close agreement with that from spectroscopic analysis. Surface gravity is derived from ionization equilibrium between Fe I and Fe II. Microturbulent velocity is estimated from individual Fe I lines by requiring the derived abundances to be independent of their EWs.

Abundance determination

The abundances of most elements lighter than Zn are determined from EWs based on the adopted stellar parameters. The upper limits of Na, Sc, Zn, Sr and Ba abundances are estimated by spectrum synthesis. Furthermore, the abundances of elements other than Fe determined from EWs analysis are also confirmed by spectrum synthesis. For lines of Sc II, Mn I and Co I, the effect of hyperfine splitting is included in the abundance determination by using the data in Kurucz's database. Both of the Na I lines at 5,889 Å and 5,895 Å are too weak for EW measurements. The upper limit of Na abundance is determined from synthesis of the Na I 5,889 Å line. We notice that the Na abundance ([Na/Fe] < −2.02) of J1010+2358 is extremely low compared with other metal-poor stars. A portion of the spectrum of a comparison star LAMOST J0626+6032 (T_{eff} = 5,863 K, log g = 3.73, [Fe/H] = −2.39, [Na/Fe] = +0.89) is shown in Extended Data Fig. 1 for comparison purposes. The spectrum of the comparison star was obtained by Subaru/HDS with the same setup. The α elements with detectable lines include Mg, Si, Ca and Ti. The Mg lines at 4,702, 5,172 and 5,183 Å were used to determine

Mg abundance. The [X/Fe] ratios for Mg, Ca and Ti are sub-solar, whereas [Si/Fe] is slightly enhanced. No absorption line of neutron-capture elements is detected in the spectrum. The upper limits of Sr and Ba are estimated from Sr II 4,077 Å and Ba II 4,554 Å, respectively. The portion of the spectrum of J1010+2358 around Ba II 4,554 Å is shown in Extended Data Fig. 1 for comparison. The carbon feature could not be detected from the molecular band of CH at 4,315 Å. Non-LTE corrections are estimated for lines of Na, Mg, Si, Ca, Cr, Mn and Fe (refs. [35,36,37,38]). The non-LTE corrections for Na I, Mg I, Si I, Mn I and Fe I are less than +0.1 dex, whereas the corrections for Ca I and Cr I are +0.16 dex and +0.21 dex, respectively. The PISN model is still the best-fitting model when the non-LTE corrections are included in the comparison between the observed abundance pattern and SN yield models.

Exclusion of CCSN models

The search algorithm STARFIT[18] compares the observed abundances of J1010+2358 against a large number of SN yield models in the literature[39,40,41,42,43,44] and determines that the best-fitting model is a 260-$M_{\odot}$ PISN with a 130-$M_{\odot}$ He core. One of the most important features of J1010+2358 is that its [Na/Fe] and [Mg/Fe] ratios are much lower than those of other metal-poor stars in the Milky Way with similar metallicities. It is impossible to reproduce such low [Na/Fe] and [Mg/Fe] abundance pattern by assuming the contributions from CCSNe because regular CCSNe cannot ever produce very low [α/Fe] ratio[45]. CCSNe with a higher explosion energy seem to be able to reduce the ratios of [Na/Fe] and [Mg/Fe] (even though it is still hard to achieve such low yields observed in J1010+2358), but they produce too high [Si/Fe] and [Co/Fe] to be consistent with the observations of J1010+2358 (refs. [15,45]). Therefore, the possibility of CCSNe as the enrichment source of J1010+2358 can be excluded.

Exclusion of SN Ia contribution

The previously known α-poor stars have been generally explained by the nucleosynthetic yields of SN Ia along with a contribution from a CCSN. The combination of a SN Ia and a normal CCSN is expected to produce a low [α/Fe] ratio because of the iron enhancement resulting from SN Ia, which is

required for explaining the low [α/Fe] ratio observed in J1010+2358. However, the low [Na/Fe], [Cr/Fe] and [Mn/Fe] ratios observed in J1010+2358 are in conflict with the expectation of the yields of SN Ia models[46,47] (Extended Data Fig. 2). It is noted that a combination of hypernovae[48] and sub-Chandrasekhar-mass SNe Ia (refs. [49,50]) may produce a lower [Mn/Fe]. But the theoretical predictions of [Si/Fe], [Ti/Fe] and [Co/Fe] from such a combination are substantially different from the observed abundances of J1010+2358. Moreover, the long interval between the two progenitors would result in enrichment from normal CCSNe, which is not consistent with the observed abundance pattern of this star.

Data availability

The data used in this analysis are available on the archive of the Japanese Virtual Observatory (http://jvo.nao.ac.jp/portal/top-page.do).

Code availability

The MOOG code for stellar line analysis and spectrum synthesis is available at https://www.as.utexas.edu/~chris/moog.html. The STARFIT tool for finding the best-fitting SN yield model is available at https://starfit.org.

References

1. Bromm, V., Coppi, P. S. & Larson, R. B. Forming the first stars in the universe: the fragmentation of primordial gas. *Astrophys. J. Lett.* **527**, L5–L8 (1999).

2. Abel, T., Bryan, G. L. & Norman, M. L. The formation and fragmentation of primordial molecular clouds. *Astrophys. J.* **540**, 39–44 (2000).

3. Susa, H., Hasegawa, K. & Tominaga, N. The mass spectrum of the first stars. *Astrophys. J.* **792**, 32 (2014).

4. Hirano, S., Hosokawa, T., Yoshida, N., Omukai, K. & Yorke, H. W. Primordial star formation under the influence of far ultraviolet radiation: 1540 cosmological haloes and the stellar mass distribution. *Mon. Not. R. Astron. Soc.* **448**, 568–587 (2015).

5. Heger, A. & Woosley, S. E. The nucleosynthetic signature of population III. *Astrophys. J.* **567**, 532–543 (2002).

6. Yong, D. et al. The most metal-poor stars. II. Chemical abundances of 190 metal-poor stars including 10 new stars with [Fe/H] ≤ -3.5. *Astrophys. J.* **762**, 26 (2013).

7. Ishigaki, M. N., Tominaga, N., Kobayashi, C. & Nomoto, K. The initial mass function of the first stars inferred from extremely metal-poor stars. *Astrophys. J.* **857**, 46 (2018).

8. Zhao, G. et al. Stellar abundance and Galactic chemical evolution through LAMOST spectroscopic survey. *Chin. J. Astron. Astrophys.* **6**, 265–280 (2006).

9. Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P. & Deng, L.-C. LAMOST spectral survey — an overview. *Res. Astron. Astrophys.* **12**, 723–734 (2012).

10. Aoki, W. et al. Four-hundred very metal-poor stars studied with LAMOST and Subaru. I. Survey design, follow-up program, and binary frequency. *Astrophys. J.* **931**, 146 (2022).

11. Li, H. et al. Four-hundred very metal-poor stars studied with LAMOST and Subaru. II. Elemental abundances. *Astrophys. J.* **931**, 147 (2022).

12. Castelli, F., Gratton, R. G. & Kurucz, R. L. Notes on the convection in the ATLAS9 model atmospheres. *Astron. Astrophys.* **318**, 841–869 (1997).

13. Sneden, C., Cowan, J. J. & Gallino, R. Neutron-capture elements in the early galaxy. *Annu. Rev. Astron. Astrophys.* **46**, 241–288 (2008).

14. Tolstoy, E., Hill, V. & Tosi, M. Star-formation histories, abundances, and kinematics of dwarf galaxies in the local group. *Annu. Rev. Astron. Astrophys.* **47**, 371–425 (2009).

15. Aoki, W., Tominaga, N., Beers, T. C., Honda, S. & Lee, Y. S. A chemical signature of first-generation very massive stars. *Science* **345**, 912–915 (2014).

16. Ivans, I. I. et al. Chemical substructure in the Milky Way halo: a new population of old stars. *Astrophys. J.* **592**, 906–934 (2003).

17. Xing, Q.-F. et al. Evidence for the accretion origin of halo stars with an extreme r-process enhancement. *Nat. Astron.* **3**, 631–635 (2019).

18. Heger, A. & Woosley, S. E. Nucleosynthesis and evolution of massive metal-free stars. *Astrophys. J.* **724**, 341–373 (2010).

19. Ohshiro, Y. et al. Discovery of a highly neutronized ejecta clump in the type Ia supernova remnant 3C 397. *Astrophys. J. Lett.* **913**, L34 (2021).

20. Weaver, T. A., Zimmerman, G. B. & Woosley, S. E. Presupernova evolution of massive stars. *Astrophys. J.* **225**, 1021–1029 (1978).

21. Heger, A., Langer, N. & Woosley, S. E. Presupernova evolution of rotating massive stars. I. Numerical method and evolution of the internal stellar structure. *Astrophys. J.* **528**, 368–396 (2000).

22. Nomoto, K., Kobayashi, C. & Tominaga, N. Nucleosynthesis in stars and the chemical enrichment of galaxies. *Annu. Rev. Astron. Astrophys.* **51**, 457–509 (2013).

23. Karlsson, T., Johnson, J. L. & Bromm, V. Uncovering the chemical signature of the first stars in the universe. *Astrophys. J.* **679**, 6–16 (2008).

24. de Bennassuti, M., Salvadori, S., Schneider, R., Valiante, R. & Omukai, K. Limits on Population III star formation with the most iron-poor stars. *Mon. Not. R. Astron. Soc.* **465**, 926–940 (2017).

25. Salvadori, S. et al. Probing the existence of very massive first stars. *Mon. Not. R. Astron. Soc.* **487**, 4261–4284 (2019).

26. Tominaga, N., Iwamoto, N. & Nomoto, K. Abundance profiling of extremely metal-poor stars and supernova properties in the early universe. *Astrophys. J.* **785**, 98 (2014).

27. Yoshii, Y. et al. Potential signature of Population III pair-instability supernova ejecta in the BLR gas of the most distant quasar at $z = 7.54$. *Astrophys. J.* **937**, 61 (2022).

28. Huang, Y. et al. Beyond spectroscopy. I. Metallicities, distances, and age estimates for over 20 million stars from SMSS DR2 and Gaia EDR3. *Astrophys. J.* **925**, 164 (2022).

29. Asplund, M., Grevesse, N., Sauval, A. J. & Scott, P. The chemical composition of the Sun. *Annu. Rev. Astron. Astrophys.* **47**, 481–522 (2009).

30. Noguchi, K. et al. High dispersion spectrograph (HDS) for the Subaru telescope. *Publ. Astron. Soc. Jpn.* **54**, 855–864 (2002).

31. Aoki, W. et al. High-resolution spectroscopy of extremely metal-poor stars from SDSS/SEGUE. I. Atmospheric parameters and chemical compositions. *Astron. J.* **145**, 13 (2013).

32. Mashonkina, L. et al. The Hamburg/ESO *R*-process enhanced star survey (HERES). V. Detailed abundance analysis of the *r*-process enhanced star HE 2327-5642. *Astron. Astrophys.* **516**, A46 (2010).

33. Sneden, C. The nitrogen abundance of the very metal-poor star HD 122563. *Astrophys. J.* **184**, 839–849 (1973).

34. Casagrande, L., Ramírez, I., Meléndez, J., Bessell, M. & Asplund, M. An absolutely calibrated T_{eff} scale from the infrared flux method. Dwarfs and subgiants. *Astron. Astrophys.* **512**, A54 (2010).

35. Lind, K., Bergemann, M. & Asplund, M. Non-LTE line formation of Fe in late-type stars — II. 1D spectroscopic stellar parameters. *Mon. Not.*

R. Astron. Soc. **427**, 50–60 (2012).

36. Zhao, G. et al. Systematic non-LTE study of the $-2.6 \leq [Fe/H] \leq 0.2$ F and G dwarfs in the solar neighborhood. II. Abundance patterns from Li to Eu. *Astrophys. J.* **833**, 225 (2016).

37. Bergemann, M. & Cescutti, G. Chromium: NLTE abundances in metal-poor stars and nucleosynthesis in the Galaxy. *Astron. Astrophys.* **522**, A9 (2010).

38. Bergemann, M. et al. Observational constraints on the origin of the elements. I. 3D NLTE formation of Mn lines in late-type stars. *Astron. Astrophys.* **631**, A80 (2019).

39. Chieffi, A. & Limongi, M. Explosive yields of massive stars from $Z = 0$ to $Z = Z_{\odot}$. *Astrophys. J.* **608**, 405–410 (2004).

40. Chieffi, A. & Limongi, M. Pre-supernova evolution of rotating solar metallicity stars in the mass range 13–120 $M_{\odot}$ and their explosive yields. *Astrophys. J.* **764**, 21 (2013).

41. Grimmett, J. J., Heger, A., Karakas, A. I. & Müller, B. Nucleosynthesis in primordial hypernovae. *Mon. Not. R. Astron. Soc.* **479**, 495–516 (2018).

42. Limongi, M. & Chieffi, A. Presupernova evolution and explosive nucleosynthesis of zero metal massive stars. *Astrophys. J. Suppl. Ser.* **199**, 38 (2012).

43. Limongi, M. & Chieffi, A. Presupernova evolution and explosive nucleosynthesis of rotating massive stars in the metallicity range $-3 \leq [Fe/H] \leq 0$. *Astrophys. J. Suppl. Ser.* **237**, 13 (2018).

44. Prantzos, N., Abia, C., Limongi, M., Chieffi, A. & Cristallo, S. Chemical evolution with rotating massive star yields – I. The solar neighbourhood and the s-process elements. *Mon. Not. R. Astron. Soc.* **476**, 3432–3459 (2018).

45. Tominaga, N., Umeda, H. & Nomoto, K. Supernova nucleosynthesis in Population III 13-50 $M_\odot$ stars and abundance patterns of extremely metal-poor stars. *Astrophys. J.* **660**, 516–540 (2007).

46. Iwamoto, K. et al. Nucleosynthesis in Chandrasekhar mass models for type IA supernovae and constraints on progenitor systems and burning-front propagation. *Astrophys. J. Suppl. Ser.* **125**, 439–462 (1999).

47. Lach, F. et al. Nucleosynthesis imprints from different Type Ia supernova explosion scenarios and implications for galactic chemical evolution. *Astron. Astrophys.* **644**, A118 (2020).

48. Kobayashi, C., Ishigaki, M. N., Tominaga, N. & Nomoto, K. The origin of low [α/Fe] ratios in extremely metal-poor stars. *Astrophys. J. Lett.* **785**, L5 (2014).

49. Badenes, C., Bravo, E., Borkowski, K. J. & Domínguez, I. Thermal X-ray emission from shocked ejecta in Type Ia supernova remnants: prospects for explosion mechanism identification. *Astrophys. J.* **593**, 358–369 (2003).

50. McWilliam, A., Piro, A. L., Badenes, C. & Bravo, E. Evidence for a sub-Chandrasekhar-mass Type Ia supernova in the Ursa Minor dwarf galaxy. *Astrophys. J.* **857**, 97 (2018).

51. Nomoto, K., Kobayashi, C. & Tominaga, N. Nucleosynthesis in stars and the chemical enrichment of galaxies. *Annu. Rev. Astron. Astrophys.* **51**, 457–509 (2013).

Acknowledgements

This study is supported by the National Natural Science Foundation of China under grant nos. 11988101, 11890694, 12222305, 12288102 and 12090040, the National Key R&D Program of China no. 2019YFA0405500, JSPS-CAS Joint Research Program and the International Partnership Program of CAS Grant no. 178GJHZ2022040GC. Q.-F.X. acknowledges the support from the Youth Innovation Promotion Association of the Chinese Academy of Sciences (no. 2020058). A.H. acknowledges software

development support from Astronomy Australia Limited's ADACS scheme (AHeger_2022B, AHeger_2023A). W.A. and M.N.I. are supported by JSPS KAKENHI grant nos. JP21H04499 and JP20H05855. Funding for LAMOST (http://www.lamost.org) has been provided by the Chinese National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. This paper includes data collected at the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. We are honoured and grateful for the opportunity to observe the Universe from Maunakea, which has cultural, historical, and natural significance in Hawaii.

Author information

Authors and Affiliations

1. CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China

   Qian-Fan Xing, Gang Zhao, Yu-Qin Chen, Hai-Ning Li & Jing-Kun Zhao

2. School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing, China

   Gang Zhao, Zheng-Wei Liu, Zhan-Wen Han & Yu-Qin Chen

3. Yunnan Observatories, Chinese Academy of Sciences, Kunming, China

   Zheng-Wei Liu & Zhan-Wen Han

4. Key Laboratory for the Structure and Evolution of Celestial Objects, Chinese Academy of Sciences, Kunming, China

   Zheng-Wei Liu & Zhan-Wen Han

5. School of Physics and Astronomy, Monash University, Clayton, Victoria, Australia

Alexander Heger

6. Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Sydney, New South Wales, Australia

   Alexander Heger

7. National Astronomical Observatory of Japan (NAOJ), Mitaka, Japan

   Wako Aoki & Miho N. Ishigaki

8. Astronomical Science Program, The Graduate University for Advanced Studies (SOKENDAI), Mitaka, Japan

   Wako Aoki & Miho N. Ishigaki

Contributions

Q.-F.X. identified the star and led the data analysis and the writing of the manuscript. G.Z. proposed and initiated this study and led the discussion. Z.-W.L., A.H. and Z.-W.H. co-led the comparison between the SN yield models and the observed abundances and contributed to the interpretation of observed results. W.A., Y.-Q.C., M.N.I., H.-N.L. and J.-K.Z. contributed to the follow-up observation and abundance analysis. All authors contributed to the discussions and the writing of the manuscript.

Corresponding author

Correspondence to Gang Zhao.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Timothy Beers and Chiaki Kobayashi for their contribution to the peer review of this work.

Additional information

Extended data

is available for this paper at https://doi.org/10.1038/s41586-023-06028-1.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Observed spectrum of J1010+2358.

The observed spectrum of J1010+2358 (filled circles: T_{eff} = 5,860 K; log g = 3.6; [Fe/H] = −2.42) are shown with a comparison star J0626+6032 (blue lines: T_{eff} = 5,863 K; log g = 3.73; [Fe/H] = −2.39). The chemical abundances ([Na/Fe] = +0.89 ± 0.08; [Mg/Fe] = +0.21 ± 0.06; [Ba/Fe] = +0.23 ± 0.06) of J0626+6032 are normal as a VMP star. The black lines are synthetic spectra with [Na/Fe] = −2.02 ± 0.3 (**a**), [Mg/Fe] = −0.66 ± 0.3 (**b**) and [Ba/Fe] = −1.37 ± 0.3 (**c**).

Extended Data Fig. 2 Comparison of observed abundance pattern and SN yield models.

The chemical abundances of J1010+2358 (red circles) compared with models of combinations of SN Ia (ref. [46]) and CCSN (refs. [15,51]) with mass of 25 $M_{\odot}$: 10% contribution from SN Ia (blue line); 30% contribution from SN Ia (green line); 50% contribution from SN Ia (purple line); 70%

contribution from SN Ia (orange line). The error bars are 1σ uncertainties of the observed abundances.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Xing, QF., Zhao, G., Liu, ZW. *et al.* A metal-poor star with abundances from a pair-instability supernova. *Nature* **618**, 712–715 (2023). https://doi.org/10.1038/s41586-023-06028-1

- Received: 13 December 2022
- Accepted: 28 March 2023
- Published: 07 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06028-1

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06028-1

| Section menu | Main menu |

- Article
- Open Access
- Published: 24 May 2023

Density-wave ordering in a unitary Fermi gas with photon-mediated interactions

- Victor Helson,
- Timo Zwettler,
- Farokh Mivehvar,
- Elvia Colella,
- Kevin Roux,
- Hideki Konishi,
- Helmut Ritsch &
- …
- Jean-Philippe Brantut

Show authors

Nature **volume 618**, pages 716–720 (2023)

- 5905 Accesses

- 93 Altmetric

- Metrics details

Abstract

A density wave (DW) is a fundamental type of long-range order in quantum matter tied to self-organization into a crystalline structure. The interplay of

DW order with superfluidity can lead to complex scenarios that pose a great challenge to theoretical analysis. In the past decades, tunable quantum Fermi gases have served as model systems for exploring the physics of strongly interacting fermions, including most notably magnetic ordering[1], pairing and superfluidity[2], and the crossover from a Bardeen–Cooper–Schrieffer superfluid to a Bose–Einstein condensate[3]. Here, we realize a Fermi gas featuring both strong, tunable contact interactions and photon-mediated, spatially structured long-range interactions in a transversely driven high-finesse optical cavity. Above a critical long-range interaction strength, DW order is stabilized in the system, which we identify via its superradiant light-scattering properties. We quantitatively measure the variation of the onset of DW order as the contact interaction is varied across the Bardeen–Cooper–Schrieffer superfluid and Bose–Einstein condensate crossover, in qualitative agreement with a mean-field theory. The atomic DW susceptibility varies over an order of magnitude upon tuning the strength and the sign of the long-range interactions below the self-ordering threshold, demonstrating independent and simultaneous control over the contact and long-range interactions. Therefore, our experimental setup provides a fully tunable and microscopically controllable platform for the experimental study of the interplay of superfluidity and DW order.

Main

Quantum gas experiments provide a unique opportunity to create complex quantum many-body systems from the bottom up by starting from a dilute gas and adding interactions in a controlled way. This was initially enabled by the precise control of the intrinsic contact interaction between atoms using Feshbach resonances[4]. Recent years have seen tremendous efforts to engineer more complex many-body systems using tailored longer-range interactions[5]. As a key extension in this direction, dipolar interactions between atoms with large permanent magnetic moment were successfully used to create supersolid phases of bosons[6]. For fermions, stronger interactions promised in polar molecules[7] or transiently realized using Rydberg dressing[8] could further lead to exotic quantum phases.

Cavity quantum electrodynamics provides a flexible platform for engineering non-local, all-to-all interactions among polarizable particles mediated by cavity photons[9,10,11]. By loading atoms inside a high-finesse cavity and driving them with a transverse pump beam in the far-detuned, dispersive regime, an effective interaction between the atoms is produced, described by an effective interaction Hamiltonian[11],

$${\hat{H}}_{{\rm{i}}{\rm{n}}{\rm{t}}}=\iint {{\rm{d}}}^{3}{\bf{r}}{{\rm{d}}}^{3}{{\bf{r}}}^{{\prime} }{\mathcal{D}}({\bf{r}},{{\bf{r}}}^{{\prime} })\hat{n}({\bf{r}})\hat{n}({{\bf{r}}}^{{\prime} }),$$
(1)

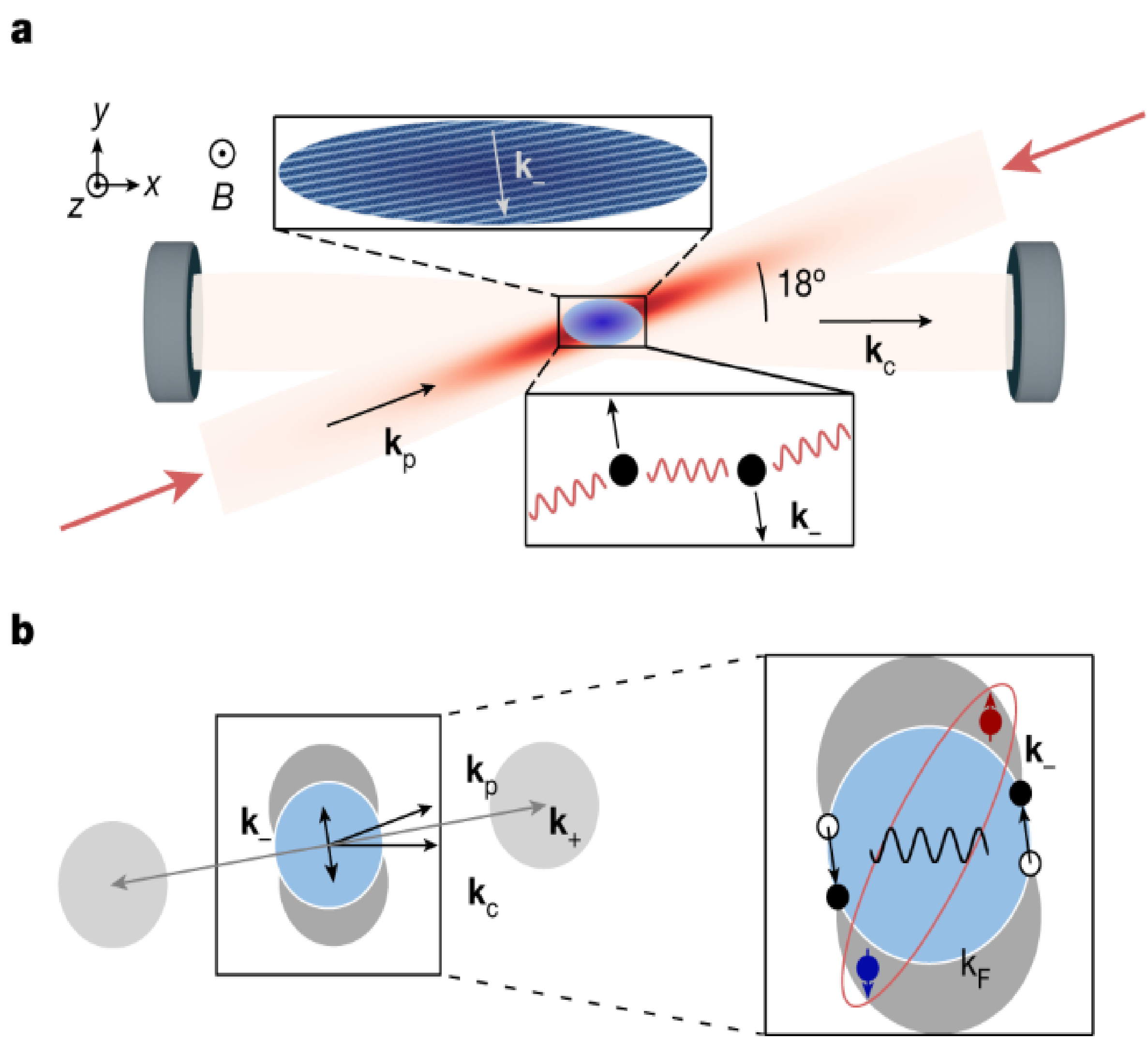
where $\hat{n}({\bf{r}})$ is the local density operator at position $\mathbf{r}$. In a single-mode cavity, this interaction has a spatially periodic, infinite-range structure of the form ${\mathcal{D}}({\bf{r}},{{\bf{r}}}^{{\prime} })={{\mathcal{D}}}_{0}\cos ({{\bf{k}}}_{{\rm{p}}}\cdot {\bf{r}})\cos ({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}})\cos ({{\bf{k}}}_{{\rm{p}}}\cdot {{\bf{r}}}^{{\prime} })\cos ({{\bf{k}}}_{{\rm{c}}}\cdot {{\bf{r}}}^{{\prime} })$, which arises from the interference of the pump and the cavity mode[12]. Here, ${{\mathcal{D}}}_{0}={U}_{0}{V}_{0}/{\varDelta }_{{\rm{c}}}$ is the interaction strength, with U_0 being the cavity potential depth per photon and V_0 being the light shift induced by the pump, proportional to the intensity of the pump laser. Δ_c is the detuning of the pump from the cavity resonance, whose sign determines the attractive or repulsive nature of the interaction (Methods). The wave vectors of pump and cavity photons are denoted by $\mathbf{k}_p$ and $\mathbf{k}_c$, respectively. Physically, the interaction Hamiltonian (equation (1)) describes the correlated recoils from the scattering of a pump photon off an atom into the cavity mode and back into the pump by a second atom.

This photon-mediated density–density interaction leads to the self-organization into a density-wave (DW) phase, as was first observed in thermal atoms[13], then in Bose–Einstein condensates (BECs)[14,15] and lattice Bose gases[16,17], and recently, in non-interacting Fermi gases[18]. In weakly interacting BECs, the DW self-ordering is a manifestation of the Dicke superradiant phase transition, and it allowed for the quantum simulation of

supersolidity[19]. By exploiting more atomic internal levels and many cavity modes, a variety of rich phenomena ranging from magnetic ordering[20,21] to dynamic gauge fields[22] and self-ordering in elastic optical lattices[23] were observed in bosonic systems. Even more intriguing phenomena ranging from threshold-less self-ordering in low dimensions to cavity-induced superconducting pairing and topological states have been predicted for fermions[24,25,26,27,28,29,30,31,32].

Here, we realize a doubly tunable Fermi gas combining simultaneously and independently the control over contact and photon-mediated long-range interactions. We explore the regime where both interactions are strong, the latter leading to DW ordering. For fermionic particles, the Pauli principle restricts the effects of interactions to the Fermi surface; thus, the resonant s-wave contact interactions yield Cooper pairing at low temperatures. By contrast, the photon-mediated interaction couples particle-hole excitations on the Fermi surface at discrete wave vectors $\mathbf{k}_{\pm} = \mathbf{k}_{c} \pm \mathbf{k}_{p}$, imposed by the pump-cavity geometry as illustrated in Fig. 1a. In our three-dimensional system, the low-energy physics is associated with scattering processes with the wave vector $\mathbf{k}_{-}$, which is smaller than the Fermi wave vector $\mathbf{k}_{F}$, leading to a broad particle-hole spectrum (in contrast to ref. [18]). This is described by the Lindhard function for free fermions, which is maximum at zero frequency for low momenta close to $\mathbf{k}_{-}$. This contrasts with large momenta, where the Pauli principle does not restrict the available phase space unless the Fermi surface is deformed[18]. We find that even in the presence of strong contact interactions, photon-mediated interactions modify the zero-frequency particle-hole susceptibility and lead to the spontaneous formation of a DW pattern above a critical strength in the attractive case.

Fig. 1: Concept of the experiment.



a, A strongly interacting Fermi gas trapped inside a high-finesse optical resonator is illuminated by a standing-wave pump laser with wave vector $\mathbf{k}_p$, polarized along the direction of the magnetic field B, which intersects the axis of the cavity mode (*x* direction) with wave vector $\mathbf{k}_c$ at an angle of 18°. The pump beam couples dispersively to atomic motion. Off-resonant scattering of pump photons by the atoms into the cavity mode and vice versa leads to an effective infinite-range interaction between atoms. Above a critical strength, the infinite-range interaction results in a superradiant phase transition to a DW-ordered state with spatial modulation at $2\pi/\mathbf{k}_{-}$. **b**, In the left panel, photon scattering from the pump into the cavity and vice versa via the atoms imparts momentum kicks $\mathbf{k}_{\pm} = \mathbf{k}_c \pm \mathbf{k}_p$ onto the latter, displacing the Fermi surface. In the right panel, since $|\mathbf{k}_{-}| < \mathbf{k}_F$, the photon-mediated

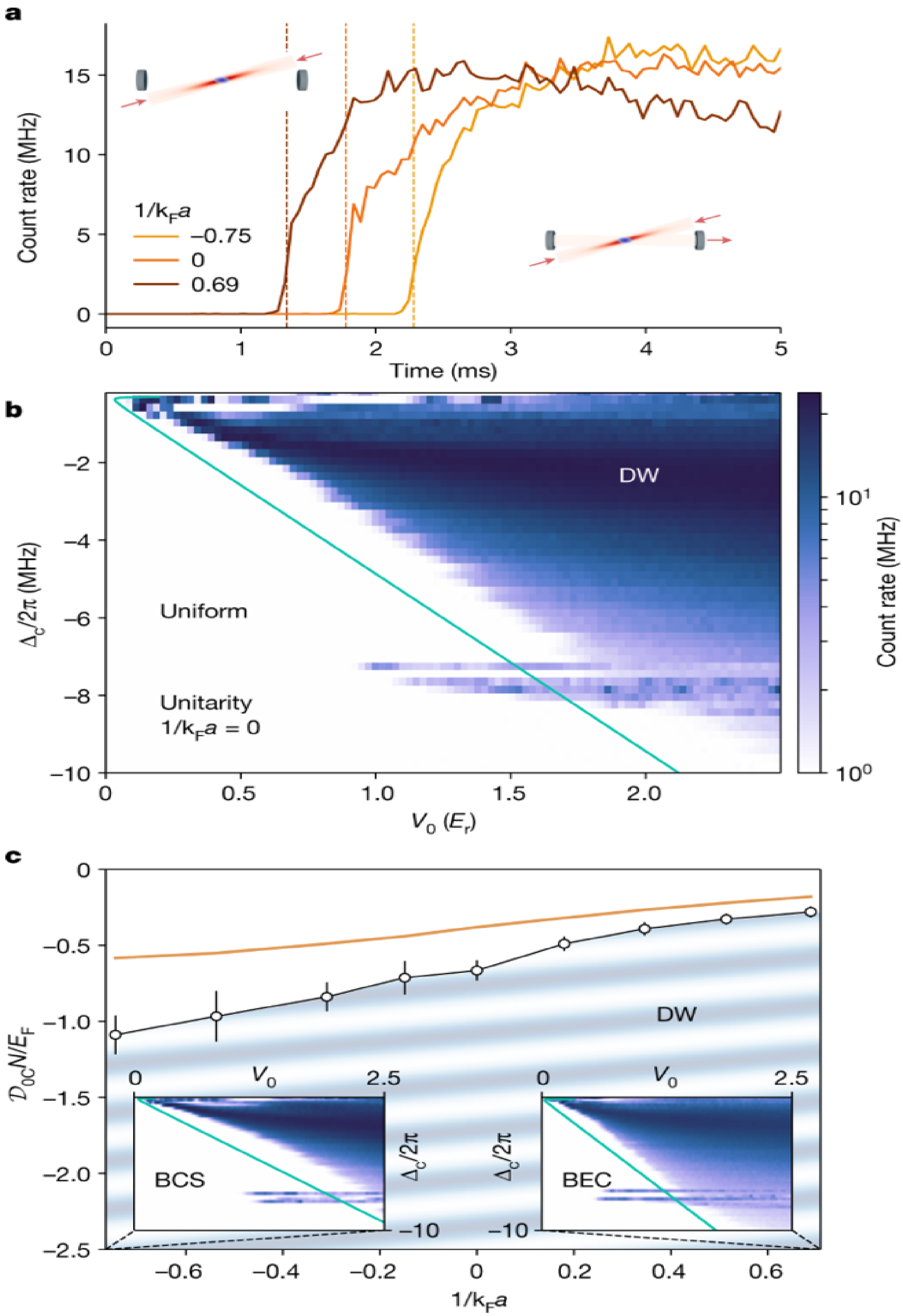
interactions induce particle-hole excitations at the Fermi surface in addition to Cooper pairing arising from the contact interactions.

In the experiment, we prepare a degenerate Fermi gas of $N = 3.5 \times 10^5$ Li atoms equally populating the two lowest hyperfine states, trapped within a mode of a high-finesse optical cavity[33,34] and in the vicinity of a broad Feshbach resonance at 832 G. We turn on the photon-mediated interactions by illuminating the cloud from the side using a retro-reflected pump beam. The pump and the neighbouring cavity resonance are detuned with respect to the atomic D_2 transition by $-2\pi \times 138.0$ GHz. There, the atoms induce a dispersive shift of the cavity resonance by $\delta_c = U_0 N/2 = -2\pi \times 280$ kHz, exceeding the cavity line width $\kappa_c = 2\pi \times 77(1)$ kHz. The pump beam intersects the cavity at an angle of 18°, such that two discrete density-fluctuation modes at momenta $\mathbf{k}_{\pm}$ are coupled to light, as illustrated in Fig. 1b. The low incidence angle results in the hierarchy $|\mathbf{k}_{-}| \ll |\mathbf{k}_{+}|$, so that only the mode at $\mathbf{k}_{-}$ contributes to the low-energy physics (Methods). We use pump-cavity detunings $|\Delta_c|/2\pi$ between 1 and 10 MHz for which $|\Delta_c| \gg |\delta_c|, \kappa_c$, and the cavity field adiabatically follows the atomic dynamics, ensuring that the system is accurately described by the Hamiltonian (equation (1)).

DW ordering

We observe DW ordering upon increasing the strength of the photon-mediated interaction above a critical threshold. Experimentally, at fixed scattering length, we linearly ramp up the pump power and monitor the intracavity photon number by recording the photon flux leaking through one of the cavity mirrors while keeping all other parameters fixed. In Fig. 2a, we show typical photon traces for different scattering lengths, as V_0 is linearly increased up to 2.5 E_r over 5 ms, with \({E}_{{\rm{r}}}={\hbar }^{2}{{\bf{k}}}_{{\rm{c}}}^{2}/2m=h\times 73.67\) kHz the recoil energy. The build-up in the cavity field above a critical pump strength V_{0C} marks the onset of DW ordering (Methods).

Fig. 2: Phase diagrams of the system.



a
Count rate (MHz)
Time (ms)
1/k_F a
−0.75
0
0.69
b
DW
Uniform
Unitarity
1/k_F a = 0
Δ_c/2π (MHz)
V_0 (E_r)
Count rate (MHz)
c
DW
BCS
BEC
V_0
Δ_c/2π
𝒟_0C N/E_F
1/k_F a

a, Photon traces recorded at fixed $\Delta_c = -2\pi \times 2$ MHz as a function of the linearly increasing pump strength V_0 for different values of the short-range interaction parameter $1/k_F a$ spanning the strongly interacting regime of the BCS–BEC crossover. Each measurement features a sharp increase of the photon count rate above a critical value of the pump strength V_{0C} (dashed vertical lines). **b**, Phase diagram of the unitary Fermi gas in the V_0–Δ_c plane, exhibiting DW self-ordering. The solid line is a theory estimate of the phase boundary (in the text). **c**, Measurement of the critical long-range interaction strength \({{\mathcal{D}}}_{0{\rm{C}}}\) as a function of the contact interaction parameter at fixed $\Delta_c = 6\delta_c$. Above the critical value, the system exhibits a modulated density, depicted by the oblique stripes. The solid line is the critical interaction strength calculated from theory. Insets display phase diagrams measured in the BCS and BEC regimes for the same parameter range as the one of **b**. Error bars represent standard deviations.

Repeating this measurement as a function of Δ_c, we construct the phase diagram of the system in the V_0–Δ_c plane, presented in Fig. 2b for the unitary gas. For small $|\Delta_c|$, the phase boundary is a straight line, corresponding to a constant ratio V_0/Δ_c, showing that the boundary is determined only by \({{\mathcal{D}}}_{0}\). For $|\Delta_c| \lesssim |\delta_c|$, we observe instabilities likely due to optomechanical effects. For $|\Delta_c| > 2\pi \times 3$ MHz, we observe a systematic deviation from the linearity, probably due to the lattice formed by the pump, changing the gas properties[35]. This single-particle effect is not captured by the effective interaction Hamiltonian (equation (1)). The structures arising at $\Delta_c \approx -2\pi \times 7$ MHz and $-2\pi \times 8$ MHz originate from the presence of high-order transverse modes of the cavity, with mode functions overlapping with the atomic density[33].

We acquire similar phase diagrams at different scattering lengths and find a transition to the DW-ordered phase for sufficiently strong pumps throughout the entire BEC and Bardeen–Cooper–Schrieffer superfluid (BCS) crossover. While the phase diagrams are qualitatively similar, with a linear phase boundary at small Δ_c, we observe a systematic shift of the DW phase boundary toward larger pump strengths as the system crosses over from the BEC to the BCS regime. In the regime $0.7\text{ MHz} < \Delta_c/2\pi < 3$ MHz, the linear

phase boundary observed at unitarity persists for all scattering lengths. This allows us to describe the DW self-ordering transition in terms of the single long-range interaction parameter $N{{\mathcal{D}}}_{0}/{E}_{{\rm{F}}}$. Figure 2c presents the phase diagram in the parameter plane of the short-range versus long-range interaction strength. We observe a smooth dependence of the phase boundary on the short-range interaction, with a systematically lower critical long-range interaction strength in the BEC side.

To understand this phase diagram, we start from the critical point ${{\mathcal{D}}}_{0{\rm{C}}}=-1/2{\chi }_{0}$, expected from the mean-field and random-phase approximations applied to the long-range interaction (Methods). Here, χ_0 is the zero-frequency susceptibility of the gas in the absence of the long-range interaction. To predict quantitatively the phase boundary in the BCS–BEC crossover, we disregard the effects of the pump lattice and the contribution of the density response at $\pm\mathbf{k}_+$ and approximate χ_0 by its long-wavelength limit, the compressibility. The latter is obtained from accurate measurements of the equation of state as a function of the scattering length[36,37]. The resulting predictions for the phase boundary are presented as solid lines in Fig. 2b,c. This simple, parameter-free theory captures very well the relative changes of the critical point across the crossover (Extended Data Fig. 3). It, however, underestimates the absolute threshold by approximately a factor of two for all short-range interaction strengths, indicating that the zero-temperature compressibility overestimates the actual susceptibility. We indeed expect that finite wave vector and finite temperature should generally decrease the susceptibility.

Susceptibility measurement

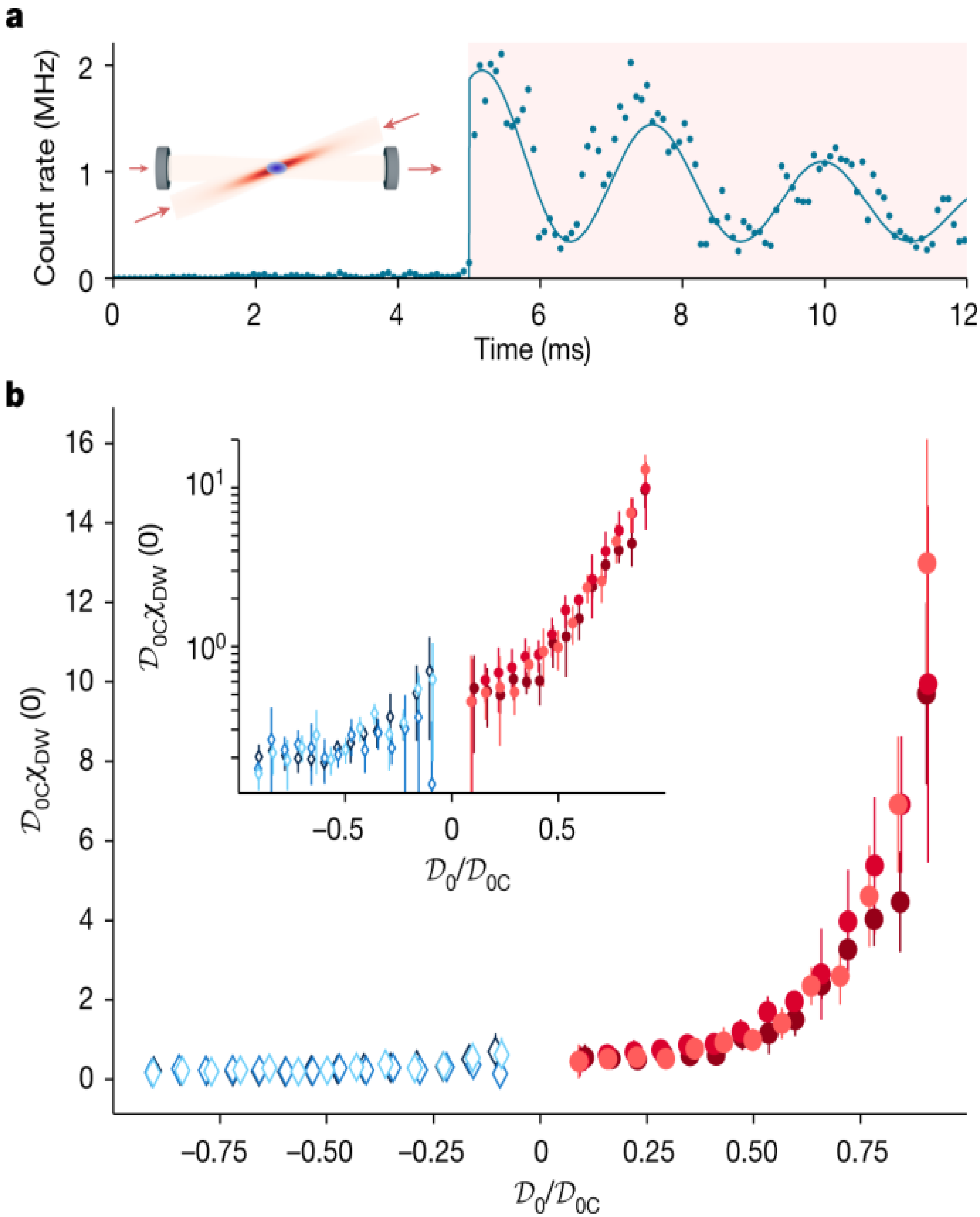
While the measurement of the cavity field allows for the identification of the onset of DW order, it does not yield information on the photon-mediated interactions below the transition. Nevertheless, the long-range interactions strongly modify properties of the gas even far below the ordering transition via virtual cavity photons. We now explore this by directly measuring the DW response function $\chi_{\rm DW}(\omega)$ as a function of the long- and short-range interaction strengths. To this end, we drive the cavity on axis using a very

weak probe laser in addition to the transverse pump[38], imposing a DW pattern at $\mathbf{k}_{\pm}$. The resulting photon-leakage rate yields χ_{DW} from the linear response theory (Methods).

In practice, the atomic response depends on the relative phase of the pump and the probe. This is intimately connected to the underlying \({{\mathbb{Z}}}_{2}\) symmetry of the model, which is broken in the ordered phase, as observed in earlier experiments on BECs[13,39]. We circumvent this issue by introducing a small detuning Δ_{p} between the pump and the probe, such that the phase winds adiabatically during the probing time, leading to slowly oscillating intracavity photon numbers. In the limit $\Delta_{\mathrm{p}} \to 0$, the amplitude of the oscillations observed in an experimental realization provides a direct measure of the zero-frequency DW response function $\chi_{\mathrm{DW}}(0)$ (Methods).

Experimentally, we first fix the long- and short-range interaction strengths by, respectively, fixing the pump power and offset magnetic field, and then, we shine the probe for 10 ms with $\Delta_{\mathrm{p}} = 2\pi \times 200$ Hz. A typical signal is shown in Fig. 3a for $\Delta_{\mathrm{c}} = -2\pi \times 2$ MHz and $V_0 = 0.75\ E_{\mathrm{r}}$, exhibiting the expected oscillations at $2\Delta_{\mathrm{p}}$ together with damping, likely due to heating resulting from the large oscillating signal. The amplitude of the initial oscillation can be directly fitted to yield the value of $\chi_{\mathrm{DW}}(0)$. For attractive photon-mediated interactions, the intracavity photon number is strongly enhanced by the presence of the atoms, as the gas coherently transfers photons from the pump to the cavity, similar to an optical parametric amplifier.

Fig. 3: Zero-frequency DW susceptibility $\chi_{\mathrm{DW}}(0)$ measurement.



a, Photon trace acquired while a weak on-axis probe beam is sent inside the cavity after the pump strength has been ramped over 5 ms to a value below the critical one. The solid line is a fit to the data (Methods), from which we extract the zero-frequency DW susceptibility $\chi_{DW}(0)$. The shaded area

highlights the interval during which the probe is on. **b**, Measured DW susceptibility as a function of the long-range interaction strength below the critical value for both attractive (red dots) and repulsive (blue diamonds) long-range interactions and for three different values of the contact interaction parameter ($1/\mathbf{k}_{\mathrm{F}}a = -0.75$, 0 and 0.69 from light to dark). The measurements were performed at constant absolute detuning $|\Delta_{\mathrm{c}} - \delta_{\mathrm{c}}| = 2\pi \times 1.7$ MHz. In the inset, the same data are displayed in logarithmic scale. Error bars represent standard deviations.

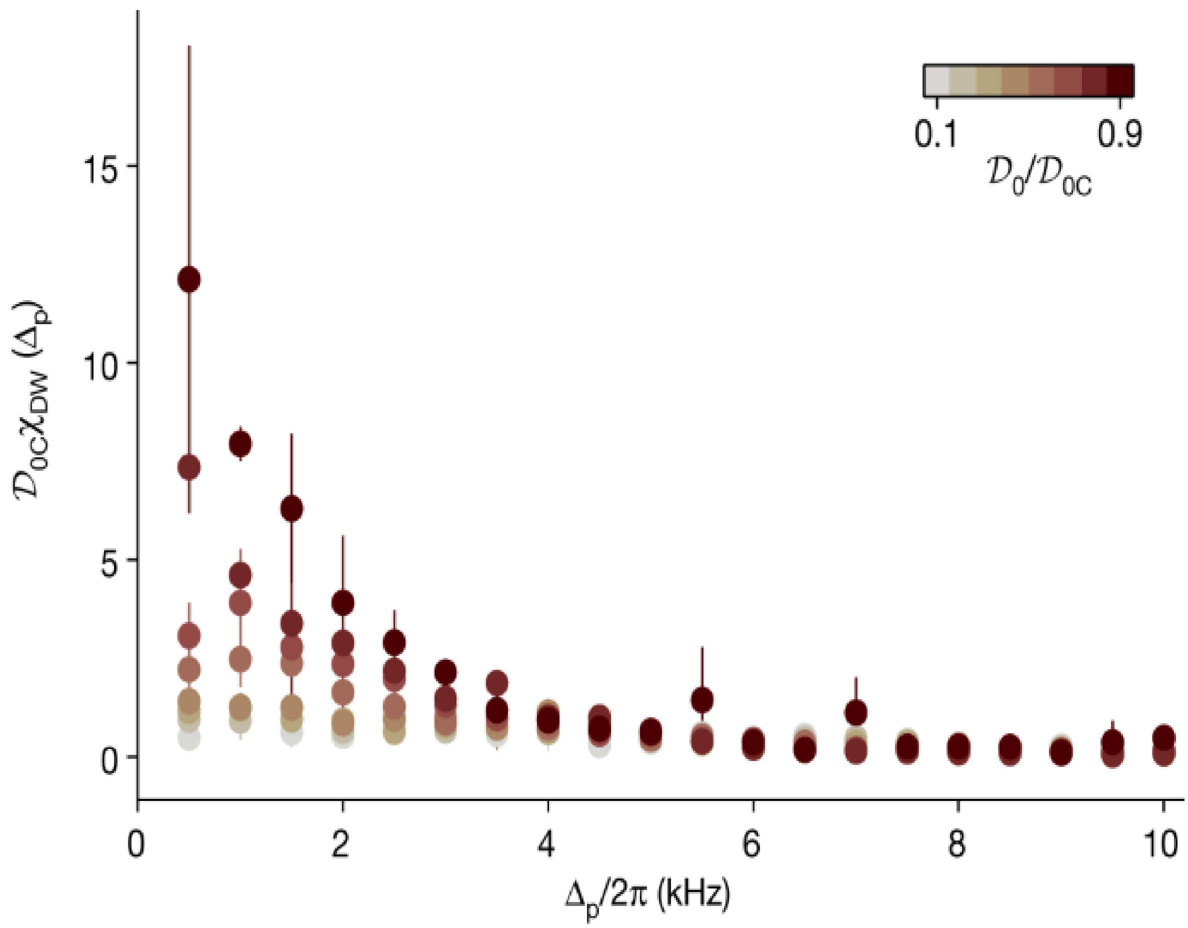
In Fig. 3b, we show the measured values of $\chi_{\mathrm{DW}}(0)$ for $\mathcal{D}_{0}$ up to $0.9\,\mathcal{D}_{0\mathrm{C}}$ at $\Delta_{\mathrm{c}} = 5\delta_{\mathrm{c}} < 0$ and $1/\mathbf{k}_{\mathrm{F}}a = -0.75$, 0 and 0.69 (red dots). We observe an increase of the susceptibility over more than one order of magnitude with increasing $\mathcal{D}_{0}$, which is the expected feature of second-order phase transitions. This was observed for self-organization and supersolid transitions in non-interacting BECs[38,40]. For repulsive photon-mediated interactions ($\Delta_{\mathrm{c}} > 0$, blue diamonds), no ordering is expected or observed, and we observe a reduction of the susceptibility by up to a factor of approximately three over the same range of $|\mathcal{D}_{0}|$. Up to normalization of $\chi_{\mathrm{DW}}(0)$ and $\mathcal{D}_{0}$ by $\mathcal{D}_{0\mathrm{C}}$, we observe that for attractive or repulsive long-range interactions, the variations of the susceptibility are identical within error bars for all scattering lengths in the BCS–BEC crossover. This highlights the versatility of our system in independently tuning the short- and long-range interactions, therefore addressing separately pairing and particle-hole channels.

The attractive (repulsive) photon-mediated interactions lower (raise) the energy cost of particle-hole excitations. For bosons with a sharp single-frequency excitation spectrum, this leads to a mode softening of the corresponding excitation mode, touching zero at the critical point[38,41]. Free fermions at low momenta, in contrast, feature a continuous, incoherent gapless particle-hole spectrum[42], such that no soft mode is expected.

We now investigate this effect for a strongly interacting Fermi gas by extending our susceptibility measurements to finite frequencies by

systematically scanning Δ_p up to $2\pi \times 10$ kHz, larger than ${\hbar }^{2}{{\bf{k}}}_{-}^{2}/2m=h\times 7.2\,$ kHz, the recoil energy associated with $\mathbf{k}_{-}$. We then extract $\chi_{DW}(\Delta_p)$ from the amplitude of the photon trace oscillations at $2\Delta_p$. For the unitary Fermi gas, the results are presented in Fig. 4 for ${{\mathcal{D}}}_{0}$ up to $0.9{{\mathcal{D}}}_{0{\rm{C}}}$, all showing that $\chi_{DW}(\Delta_p)$ monotonically decreases with frequency Δ_p. The low-frequency susceptibility increases upon approaching the transition, while the higher-frequencies parts of the spectrum remain unchanged. We observe such a behaviour for all accessible scattering lengths in the BCS–BEC crossover. This contrasts with the mode softening observed with weakly interacting BECs. While this would be expected in our geometry for free fermions, due to the broad particle-hole spectrum, it is surprising that this feature is also present for the unitary Fermi gas, which is known to also display a phonon spectrum at low momentum[43,44]. This might be due to the strongly interacting nature of the system leading to the damping of the excitations but could also originate from the combination of finite temperature and trap averaging.

Fig. 4: Measurement of the DW response $\chi_{DW}(\Delta_p)$ of a unitary Fermi gas as a function of Δ_p and for values of ${\boldsymbol{\mathcal{D}}}_{{\bf{0}}}/{\boldsymbol{\mathcal{D}}}_{{\bf{0}}{\rm{C}}}$ between 0.1 and 0.9.



The absence of structure at finite frequency confirms the absence of mode softening. The data are taken for $\Delta_c = -2\pi \times 2$ MHz. Error bars represent standard deviations.

Discussion

We operate with atoms in the deeply degenerate regime with temperatures on the order of $T = 0.08\ T_{Fh}$, with T_{Fh} the Fermi temperature calculated for a harmonic trap, where for all interaction strengths, the system is superfluid in the absence of the photon-mediated interactions. For a wide range of the short-range interaction strength, the system enters the DW-ordered phase upon increasing the photon-mediated interaction strength and returns to the superfluid phase when the long-range interaction is ramped back to zero, with limited heating (Extended Data Fig. 1). However, this leaves open the fascinating question of whether the system remains paired and superfluid in the presence of strong long-range interactions and in the DW-ordered state.

Compared with condensed-matter systems showing an interplay of charge DW and superfluidity[45], our system has a fully controllable microscopic Hamiltonian. The photon-induced DW order shares similarities with type II charge-DW compounds[46], with cavity photons playing the role of phonons in real materials. In this context, the real-time weakly destructive measurement channel through the cavity field opens the possibility of gaining insight into the interplay of structural effects and strong interactions in complex quantum materials.

Our platform complements ongoing research in the field of cavity-coupled strongly correlated materials, where the cavity photons couple to the kinetic energy of charges through the Peierls phase[47,48] or indirectly via interband transitions or collective modes. Interestingly, a direct two-photon density coupling similar to ours has been predicted for side-pumped two-dimensional materials, originating from diamagnetic interactions between charges and light and leading to enhanced superconductivity[49].

Natural extensions of our experiment include the use of several pumping frequencies addressing multiple cavity modes, providing further control over the long-range interaction potential[12,23], and the study of retardation effects due to our cavity line width being comparable with the photon recoil energy at k_c (ref. [15]). A fascinating perspective is to operate the pump in the vicinity of a photo-association transition[50], offering the possibility to induce long-range pair–pair interactions.

Methods

Experimental procedure

We produce a strongly interacting Fermi gas of ^6Li following the method described in refs. [33,34]. This procedure produces deeply degenerate, balanced mixtures of the two lowest hyperfine states trapped in a crossed dipole trap elongated along the cavity axis, formed by two Gaussian laser beams with waists of 33 μm intersecting each other with an angle of 36°.

Thermometry is performed by releasing the cloud into a hybrid trap, formed by one of the arms of the dipole trap and the residual curvature of the magnetic field[34]. An in situ absorption image is then taken with a light intensity optimized for the signal-to-noise ratio, and the density profile is obtained from the image using finite-saturation corrections. The reduced temperature in this trap is deduced from the shape of the cloud at unitarity. This yields a $T/T_{\rm Fh}$ with ${T}_{{\rm{Fh}}}=\hbar \bar{\omega }{(3N)}^{1/3}$, with N the total number of atoms and $\bar{\omega }={({\omega }_{x}{\omega }_{y}{\omega }_{z})}^{1/3}=2{\rm{\pi }}\times 106$ Hz is the geometric mean of the oscillation frequencies in the hybrid trap. This provides us with an upper bound of the degree of degeneracy in the crossed dipole trap.

The hybrid trap is harmonic and allows for both precise thermometry and calibration of each beam geometry. To reach the lowest temperatures, we found out that the crossed dipole trap operates in a regime where the anharmonicity is too strong to allow for harmonic approximation. For the purpose of evaluating the theoretical phase boundary, we instead use the full crossed-Gaussian beam trap shape deduced from trap frequencies measured in each beam separately. We then deduce the density distribution using the zero-temperature equation of state in the BEC–BCS crossover[36,37].

The pump beam is linearly polarized along the magnetic-field direction, and we estimate its waist to be 120 μm, much larger than the Thomas–Fermi radii of the cloud. We calibrate the depth of the pump lattice using Kapitza–Dirac diffraction on a molecular BEC at $B = 695$ G (ref. [51]). The photons leaking from one of the cavity mirrors are detected using a single-photon counting module with an efficiency of approximately 3% (ref. [52]).

Heating due to the side pumping

We estimate the heating due to the pump by measuring the temperature of the cloud after linearly ramping up the pump lattice depth to varying end values at a constant rate and then, ramping it back to zero with the same rate. With increasing pump power, we observe a monotonically increasing temperature of the cloud shown in Extended Data Fig. 1. Interestingly, temperature shows no particular feature when the pump power reaches and

exceeds the DW-ordering threshold. At the critical point, we measure a temperature of $T = 0.12(2)T_{\rm Fh}$, an increase by a factor of 50% compared with the initial one. Heating is sufficient to heat the cloud above the superfluid critical temperature of $0.21T_{\rm Fh}$ (ref. [53]) for a strength of the long-range interactions exceeding $2{{\mathcal{D}}}_{0{\rm{C}}}$, deep in the ordered phase. By extracting the atom number from the density profiles, we verify that the losses display the same trend with varying pump strength.

Theoretical model

The Fermi gas is coupled to a single standing-wave mode designated by the operator $\hat{a}$ of the cavity with the single atom-photon coupling strength $g({\bf{r}})={g}_{0}\cos ({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}})$, where $\mathbf{k}_{\rm c} = |\mathbf{k}_{\rm c}|\mathbf{e}_{\rm x} = k_{\rm c}\mathbf{e}_{\rm x}$ is the cavity wave vector. The atomic cloud is also transversely pumped by an incident, back-reflected pump laser with the wave vector $\mathbf{k}_{\rm p}$, where $k_{\rm p} = |\boldsymbol{k}_{\rm p}| \simeq k_{\rm c}$ and frequency $\omega_{\rm p} = ck_{\rm p}$. In the dispersive regime, the atoms experience an effective lattice potential[11], identical for the two hyperfine components of the gas:

$$\begin{array}{ll}{\hat{V}}_{{\rm{latt}}}({\bf{r}}) & =\,{V}_{0}{\cos }^{2}({{\bf{k}}}_{{\rm{p}}}\cdot {\bf{r}})+{U}_{0}{\hat{a}}^{\dagger }\hat{a}{\cos }^{2}({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}})\\ & +{\eta }_{0}(\hat{a}+{\hat{a}}^{\dagger })\cos ({{\bf{k}}}_{{\rm{p}}}\cdot {\bf{r}})\cos ({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}}),\end{array}$$
(2)

where ${\eta }_{0}=\sqrt{{V}_{0}{U}_{0}}$. This potential is added to the external trap potential ${V}_{{\rm{tr}}}({\bf{r}})$.

In the frame rotating at the pump-laser frequency, the system is described by the Hamiltonian (we set $\hbar = 1$ throughout this section),

$$\begin{array}{c}\hat{H}=-{\Delta }_{{\rm{c}}}{\hat{a}}^{\dagger }\hat{a}+\beta [\hat{a}{e}^{-{\rm{i}}({\Delta }_{{\rm{p}}}{\rm{t}}-{\varphi }_{0})}+{\rm{H}}\,.\,{\rm{c}}.]\\ \,+\sum _{{\sigma }=\downarrow ,\uparrow }\times \int {\rm{d}}{\bf{r}}{\hat{\Psi }}_{{\sigma }}^{\dagger }({\bf{r}})\,\left[-\frac{{{\rm{\nabla }}}^{2}}

{2m}-{\mu }_{{\sigma }}+{V}_{{\rm{t}}{\rm{r}}}({\bf{r}})+{\hat{V}}_{{\rm{l}}{\rm{a}}{\rm{t}}{\rm{t}}}({\bf{r}})\right]{\hat{\Psi }}_{{\sigma }}({\bf{r}})\\ \,+\int {\rm{d}}{\bf{r}}{\rm{d}}{{\bf{r}}}^{{\prime} }{\hat{\Psi }}_{\uparrow }^{\dagger }({\bf{r}}){\hat{\Psi }}_{\downarrow }^{\dagger }({{\bf{r}}}^{{\prime} }){V}_{{\rm{s}}{\rm{r}}}({\bf{r}}-{{\bf{r}}}^{{\prime} }){\hat{\Psi }}_{\downarrow }({{\bf{r}}}^{{\prime} }){\hat{\Psi }}_{\uparrow }({\bf{r}}),\end{array}$$
(3)

where the first term is the free cavity Hamiltonian with the pump-cavity detuning $\Delta_{c} = \omega_{p} - \omega_{c}$, \({\hat{\Psi }}_{\sigma }({\bf{r}})\) is the fermionic annihilation-field operator for spin $\sigma = \{\downarrow, \uparrow\}$, μ_{σ} is the chemical potential, and \({V}_{{\rm{sr}}}({\bf{r}}-{{\bf{r}}}^{{\prime} })\) is a pseudopotential yielding the s-wave scattering length a between two atoms[54]. For later use, we have also included an on-axis probe with strength β, the pump-probe detuning $\Delta_{p} = \omega_{p} - \omega_{probe}$ and an initial phase ϕ_{0}. In the experiment, $\beta = 0$, except for the purpose of measuring the DW response function $\chi_{DW}(\omega)$ (see the main text and Linear response theory and the DW response function χDW(ω)).

The Hamiltonian equation (3) can be recast in the form

$$\begin{array}{ll}\hat{H} & =\,{\hat{H}}_{{\rm{at}}}-{\hat{\widetilde{\Delta }}}_{{\rm{c}}}{\hat{a}}^{\dagger }\hat{a}+{\eta }_{0}(\hat{a}+{\hat{a}}^{\dagger })\hat{\Theta }\\ & +\beta \left[\hat{a}{e}^{-{\rm{i}}({\Delta }_{{\rm{p}}}{\rm{t}}-{\varphi }_{0})}+{\rm{H.c.}}\right],\end{array}$$
(4)

where \({\hat{H}}_{{\rm{at}}}\) is the Hamiltonian of an interacting, trapped two-component Fermi gas with a classical lattice potential \({V}_{{\rm{p}}}({\bf{r}})={V}_{0}{\cos }^{2}({{\bf{k}}}_{{\rm{p}}}\cdot {\bf{r}})\) formed by the pump. Here, \({\hat{\widetilde{\Delta }}}_{{\rm{c}}}={\Delta }_{{\rm{c}}}-{\hat{\delta }}_{{\rm{c}}}={\Delta }_{{\rm{c}}}-{U}_{0}\int \,d{\bf{r}}\,{\cos }^{2}({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}})\hat{n}

({\bf{r}})\), with \(\hat{n}({\bf{r}})={\sum }_{\sigma }{\hat{n}}_{\sigma }({\bf{r}})={\sum }_{\sigma }{\hat{\Psi }}_{\sigma }^{\dagger }({\bf{r}}){\hat{\Psi }}_{\sigma }({\bf{r}})\) being the total density operator, is the dispersively shifted pump-cavity detuning and

$$\begin{array}{c}\hat{\Theta }=\int {\rm{d}}{\bf{r}}\,\cos ({{\bf{k}}}_{{\rm{p}}}\cdot {\bf{r}})\cos ({{\bf{k}}}_{{\rm{c}}}\cdot {\bf{r}})\hat{n}({\bf{r}})\\ \,=\,\frac{1}{4}({\hat{n}}_{{{\bf{k}}}_{+}}+{\hat{n}}_{-{{\bf{k}}}_{+}}+{\hat{n}}_{{{\bf{k}}}_{-}}+{\hat{n}}_{-{{\bf{k}}}_{-}}),\end{array}$$
(5)

with \({\hat{n}}_{{\bf{q}}}=\int \,d{\bf{r}}\hat{n}({\bf{r}}){e}^{{\rm{i}}{\bf{q}}\cdot {\bf{r}}}\) being the Fourier component of the total density operator, is the atomic DW operator describing the modulation of the atomic density at wave vectors $\mathbf{k}_{\pm} = \mathbf{k}_{\mathrm{p}} \pm \mathbf{k}_{\mathrm{c}}$.

In the Hamiltonian equation (4) describing our experiment, Δ_{c} is much larger than all other energy scales (including the dispersive shift \({\delta }_{{\rm{c}}}=\langle {\hat{\delta }}_{{\rm{c}}}\rangle \), so that \({\widetilde{\Delta }}_{{\rm{c}}}=\langle {\hat{\widetilde{\Delta }}}_{{\rm{c}}}\rangle \simeq {\Delta }_{{\rm{c}}}\)), so that the cavity-field dynamics is very fast and follows the atomic dynamics. The steady-state cavity-field operator can, therefore, be obtained through the Heisenberg equation of motion, yielding

$$\hat{a}=\frac{1}{{\Delta }_{{\rm{c}}}+i{\kappa }_{{\rm{c}}}}\left[[{\eta }_{0}\hat{\Theta }+\beta {e}^{{\rm{i}}({\Delta }_{{\rm{p}}}{\rm{t}}-{\varphi }_{0})}]\right].$$
(6)

Substituting the steady-state cavity-field operator (6) in the Hamiltonian (4) and ignoring a constant term yields an effective, atom-only description of the system (up to the inverse square of the detuning of the pump laser with respect to the atomic transition)[11]:

$${\hat{H}}_{{\rm{eff-at}}}={\hat{H}}_{{\rm{at}}}+{{\mathcal{D}}}_{0}{\hat{\Theta }}^{2}+\frac{2\beta }{{\eta }_{0}}

$${{\mathcal{D}}}_{0}\hat{\Theta }\cos ({\Delta }_{{\rm{p}}}t-{\varphi }_{0}),$$
(7)

where ${{\mathcal{D}}}_{0}={\Delta }_{{\rm{c}}}{\eta }_{0}^{2}\,/({\Delta }_{{\rm{c}}}^{2}+{\kappa }_{{\rm{c}}}^{2})\simeq {\eta }_{0}^{2}\,/{\Delta }_{{\rm{c}}}$ is the strength of the cavity-mediated long-range density–density interaction. In the last equality, we asserted $\kappa_c \ll \Delta_c$, as realized in the experiment. The last term in equation (7) is the driving of the Fermi gas due to the interference between the pump and the on-axis probe.

Theoretical phase boundary

We identify the critical pump threshold ${\eta }_{0{\rm{C}}}=\sqrt{{V}_{0{\rm{C}}}{U}_{0}}$ that separates the superradiant phase from the normal state through perturbation theory[55] by integrating out the atomic degrees of freedom and expanding the resultant free energy in powers of the order parameter $\hat{\Theta }$. Up to second order in the order parameter, we obtain the free energy as in Landau theory,

$$F\approx ({\eta }_{0{\rm{C}}}^{2}-{\eta }_{0}^{2}){\hat{\Theta }}^{2}+{\rm{O}}({\hat{\Theta }}^{4}),$$
(8)

where ${\eta }_{0{\rm{C}}}^{2}=-\,({\Delta }_{{\rm{c}}}^{2}+{\kappa }_{{\rm{c}}}^{2})/2{\Delta }_{{\rm{c}}}\,{\chi }_{0}\simeq -{\Delta }_{{\rm{c}}}/2{\chi }_{0}$. This corresponds to the critical long-range interaction strength ${{\mathcal{D}}}_{0{\rm{C}}}=-\,1/2{\chi }_{0}$, where χ_0 denotes the atomic susceptibility representing the response of the interacting Fermi gas to density perturbations at the wave vectors $\mathbf{k}_{\pm}$ in the absence of the pump and cavity lattices:

$${\chi }_{0}=\frac{1}{16}\sum _{{\bf{q}}=\pm {{\bf{k}}}_{\pm }}{\chi }_{0}^{{\rm{R}}}({\bf{q}}).$$
(9)

Here, ${\chi }_{0}^{{\rm{R}}}({\bf{q}})$ is the retarded density–density response function at zero frequency and wave vector $\mathbf{q}$, calculated at a fixed finite scattering length. It coincides with the Lindhard function for a non-interacting Fermi gas.

To compare with the experiment, we first note that the short-wavelength contributions to χ_0 at $\pm\mathbf{k}_+$ are negligible compared with the low momentum one. Indeed, for $|\mathbf{k}_+| \gg \mathbf{k}_F$, the density response can be evaluated in the BCS–BEC crossover using operator product expansion[52], yielding to lowest order ${\chi }_{0}^{{\rm{R}}}({{\bf{k}}}_{+})\approx 2N/{{\epsilon }}_{{{\bf{k}}}_{+}}$ with ${{\epsilon }}_{{{\bf{k}}}_{+}}={\hbar }^{2}{{\bf{k}}}_{+}^{2}/2m$. Throughout the BCS–BEC crossover, the ratio ${\chi }_{0}^{{\rm{R}}}({{\bf{k}}}_{-})/{\chi }_{0}^{{\rm{R}}}({{\bf{k}}}_{+})$ is the smallest in the far-BCS regime and bounded from below by $3{{\epsilon }}_{{{\bf{k}}}_{+}}/4{E}_{{\rm{F}}}$, which is approximately 12 for our parameters.

We then evaluate the long-wavelength contributions ${\chi }_{0}^{{\rm{R}}}({\bf{q}}=\pm {{\bf{k}}}_{-})$. For $\mathbf{q} \to 0$, the compressibility sum rule gives ${\chi }_{0}^{{\rm{R}}}(0)=\partial n/\partial \mu ={n}^{2}\kappa$, with κ being the compressibility. For low but finite $\mathbf{q} = \pm\, \mathbf{k}_-$, hydrodynamics is expected to provide a good description of the density response, which suggests that ${\chi }_{0}^{{\rm{R}}}({\bf{q}})$ is essentially independent of momentum[56]. We therefore use the compressibility κ inferred from the thermodynamic equation of state as an estimate of ${\chi }_{0}^{{\rm{R}}}(\pm {{\bf{k}}}_{-})$ in the BCS–BEC crossover. The equation of state of a homogeneous Fermi gas has been measured accurately as a function of the contact interaction strength[36,37]. We use the interpolation formula for the universal thermodynamic functions provided in ref. [36] to deduce the compressibility of the homogeneous Fermi gas. We then use the local density approximation to perform trap averaging and to relate it to the Fermi energy E_F at the centre of the trap.

Linear response theory and the DW response function $\chi_{\mathrm{DW}}(\omega)$

We now turn our attention to the last term of equation (7) arising from the on-axis pumping of the cavity mode. We calculate the response of the DW order operator to first order using the Kubo formula,

$$\begin{array}{c}\langle \hat{\Theta }(t)\rangle ={\langle \hat{\Theta }\rangle }_{0}\\ \,\,\,+\frac{2\beta {{\mathcal{D}}}_{0}}{{\eta }_{0}}{\int }_{-{\rm{\infty }}}^{{\rm{\infty }}}{\rm{d}}{t}^{{\prime} }{\chi }_{{\rm{D}}{\rm{W}}}(t-{t}^{{\prime} })\cos ({\Delta }_{{\rm{p}}}{t}^{{\prime} }-{\varphi }_{0}),\end{array}$$
(10)

where the DW response function ${\chi }_{{\rm{DW}}}(t-{t}^{{\prime} })$ is given by

$$\begin{array}{r}{\chi }_{{\rm{DW}}}(t-{t}^{{\prime} })=-\,i\theta (t-{t}^{{\prime} }){\left\langle [\hat{\Theta }(t),\hat{\Theta }({t}^{{\prime} })]\right\rangle }_{0}.\end{array}$$
(11)

Here, $\theta(t)$ is the unit step function and $\langle ...\rangle_0$ implies averaging with $\beta = 0$.

Introducing the Fourier transform ${\chi }_{{\rm{DW}}}({\Delta }_{{\rm{p}}})={\int }_{-\infty }^{\infty }d\tau {\chi }_{{\rm{DW}}}(\tau){e}^{-{\rm{i}}{\Delta }_{{\rm{p}}}\tau }$ and noting that ${\chi }_{{\rm{DW}}}({\Delta }_{{\rm{p}}})={\chi }_{{\rm{DW}}}^{* }(-{\Delta }_{{\rm{p}}})$, equation (10) can be recast as

$$\begin{array}{r}\delta \langle \hat{\Theta }(t)\rangle =\frac{2\beta {{\mathcal{D}}}_{0}}{{\eta }_{0}}{\rm{Re}}\,\left[{\chi }_{{\rm{DW}}}({\Delta }_{{\rm{p}}}){e}^{{\rm{i}}({\Delta }_{{\rm{p}}}{\rm{t}}-{\phi }_{0})}\right],\end{array}$$
(12)

where $\delta \langle \hat{\Theta }(t)\rangle \equiv \langle \hat{\Theta }(t)\rangle -{\langle \hat{\Theta }\rangle }_{0}$. In the low-frequency limit $\Delta_{\rm p} \ll c_{\rm s}|\mathbf{k}_|$, where $c_{\rm s}$ is the speed of sound, the dynamical response function is purely real and ${\chi }_{{\rm{DW}}}({\Delta }_{{\rm{p}}})\simeq {\chi }_{{\rm{DW}}}(0)+O({({\Delta

}_{{\rm{p}}}/{c}_{{\rm{s}}}| {{\bf{k}}}_{-}|)}^{2})\) such that we obtain

$$\delta \left\langle \hat{\Theta }(t)\right\rangle \simeq \frac{2\beta {{\mathcal{D}}}_{0}}{{\eta }_{0}}{\chi }_{{\rm{DW}}}(0)\cos ({\Delta }_{{\rm{p}}}t-{\varphi }_{0}).$$
(13)

Below the superradiant threshold, $ {\langle \hat{\Theta }\rangle }_{0}=0 $, and the intracavity photon signal to first order then reads

$$\left\langle {\hat{a}}^{\dagger }\hat{a}\right\rangle =\frac{{\beta }^{2}}{{\Delta }_{{\rm{c}}}^{2}}\left[1+4{{\mathcal{D}}}_{0}\,{\chi }_{{\rm{DW}}}(0){\cos }^{2}({\Delta }_{{\rm{p}}}t-{\varphi }_{0})\right],$$
(14)

relating the oscillation in the intracavity photon number to the DW susceptibility $\chi_{\rm DW}(0)$.

Data analysis

The value of the critical pump depth V_{0C} at which the system undergoes the phase transition is inferred from photons leaking out of the cavity while the pump depth is increased. For a single realization of the experiment, we construct the histogram of arrival times of photons on the detector as a function of the pump depth, which increases linearly with time. Then, V_{0C} is determined from the point at which the slope of the reconstructed photon trace is the highest, obtained from taking its numerical derivative.

We extract $\chi_{\rm DW}(0)$ from a fit of measured photon traces to the model described by equation (14). We account for the amplitude decay of the oscillation through the addition of a factor $e^{-t/\tau}$ to the oscillatory term of the model. This may in particular capture heating and atomic losses during the measurement. Interestingly, the damping factor $1/\tau$ of the measured response features a continuous increase as the pump power approaches the threshold, as shown in Extended Data Fig. 2. The phase offset ϕ_0 is distributed

uniformly over $[0, \pi]$ for different realizations, as expected for a random relative phase between the pump and the probe. We verified that for all values of pump power, the fitted amplitude of the response varies linearly with the probe power, validating the linear response hypothesis underlying the fit.

Data availability

All data files are available from the corresponding author upon request. Accompanying data, including those for figures, are available from the Zenodo repository (https://zenodo.org/record/7733304).

References

1. Gross, C. & Bloch, I. Quantum simulations with ultracold atoms in optical lattices. *Science* **357**, 995 (2017).

2. Ketterle, W. & Zwierlein, M. W. Making, probing and understanding ultracold Fermi gases. *Nuovo Cimento Rivista Serie* **31**, 247 (2008).

3. Strinati, G. C. et al. The BCS–BEC crossover: from ultra-cold fermi gases to nuclear systems. *Phys. Rep.* **738**, 1–76 (2018).

4. Chin, C., Grimm, R., Julienne, P. & Tiesinga, E. Feshbach resonances in ultracold gases. *Rev. Mod. Phys.* **82**, 1225–1286 (2010).

5. Defenu, N. et al. Long-range interacting quantum systems. Preprint at https://arxiv.org/abs/2109.01063 (2021).

6. Chomaz, L. et al. Dipolar physics: a review of experiments with magnetic quantum gases. *Rep. Prog. Phys.* **86**, 026401 (2022).

7. Moses, S. A., Covey, J. P., Miecnikowski, M. T., Jin, D. S. & Ye, J. New frontiers for quantum gases of polar molecules. *Nat. Phys.* **13**, 13–20 (2017).

8. Guardado-Sanchez, E. et al. Quench dynamics of a fermi gas with strong nonlocal interactions. *Phys. Rev. X* **11**, 021036 (2021).

9. Münstermann, P., Fischer, T., Maunz, P., Pinkse, P. W. H. & Rempe, G. Observation of cavity-mediated long-range light forces between strongly coupled atoms. *Phys. Rev. Lett.* **84**, 4068–4071 (2000).

10. Ritsch, H., Domokos, P., Brennecke, F. & Esslinger, T. Cold atoms in cavity-generated dynamical optical potentials. *Rev. Mod. Phys.* **85**, 553–601 (2013).

11. Mivehvar, F., Piazza, F., Donner, T. & Ritsch, H. Cavity QED with quantum gases: new paradigms in many-body physics. *Adv. Phys.* **70**, 1–153 (2021).

12. Vaidya, V. D. et al. Tunable-range, photon-mediated atomic interactions in multimode cavity QED. *Phys. Rev. X* **8**, 011002 (2018).

13. Black, A. T., Chan, H. W. & Vuletić, V. Observation of collective friction forces due to spatial self-organization of atoms: from Rayleigh to Bragg scattering. *Phys. Rev. Lett.* **91**, 203001 (2003).

14. Baumann, K., Guerlin, C., Brennecke, F. & Esslinger, T. Dicke quantum phase transition with a superfluid gas in an optical cavity. *Nature* **464**, 1301–1306 (2010).

15. Klinder, J., Keßler, H., Wolke, M., Mathey, L. & Hemmerich, A. Dynamical phase transition in the open Dicke model. *Proc. Natl Acad. Sci. USA* **112**, 3290–3295 (2015).

16. Klinder, J., Keßler, H., Bakhtiari, M. R., Thorwart, M. & Hemmerich, A. Observation of a superradiant Mott insulator in the Dicke-Hubbard model. *Phys. Rev. Lett.* **115**, 230403 (2015).

17. Landig, R. et al. Quantum phases from competing short- and long-range interactions in an optical lattice. *Nature* **532**, 476–479 (2016).

18. Zhang, X. et al. Observation of a superradiant quantum phase transition in an intracavity degenerate fermi gas. *Science* **373**, 1359–1362 (2021).

19. Léonard, J., Morales, A., Zupancic, P., Esslinger, T. & Donner, T. Supersolid formation in a quantum gas breaking a continuous translational symmetry. *Nature* **543**, 87–90 (2017).

20. Landini, M. et al. Formation of a spin texture in a quantum gas coupled to a cavity. *Phys. Rev. Lett.* **120**, 053603 (2018).

21. Kroeze, R. M., Guo, Y., Vaidya, V. D., Keeling, J. & Lev, B. L. Spinor self-ordering of a quantum gas in a cavity. *Phys. Rev. Lett.* **121**, 163601 (2018).

22. Kroeze, R. M., Guo, Y. & Lev, B. L. Dynamical spin-orbit coupling of a quantum gas. *Phys. Rev. Lett.* **123**, 160404 (2019).

23. Guo, Y. et al. An optical lattice with sound. *Nature* **599**, 211–215 (2021).

24. Keeling, J., Bhaseen, M. J. & Simons, B. D. Fermionic superradiance in a transversely pumped optical cavity. *Phys. Rev. Lett.* **112**, 143002 (2014).

25. Piazza, F. & Strack, P. Umklapp superradiance with a collisionless quantum degenerate fermi gas. *Phys. Rev. Lett.* **112**, 143003 (2014).

26. Chen, Y., Yu, Z. & Zhai, H. Superradiance of degenerate fermi gases in a cavity. *Phys. Rev. Lett.* **112**, 143004 (2014).

27. Yang, S., Al-Amri, M. & Zubairy, M. S. Dicke quantum phase transition with a degenerate fermi gas in an optical cavity. *J. Phys. B At. Mol. Opt. Phys.* **47**, 135503 (2014).

28. Chen, Y., Zhai, H. & Yu, Z. Superradiant phase transition of Fermi gases in a cavity across a Feshbach resonance. *Phys. Rev. A* **91**, 021602 (2015).

29. Kollath, C., Sheikhan, A., Wolff, S. & Brennecke, F. Ultracold fermions in a cavity-induced artificial magnetic field. *Phys. Rev. Lett.* **116**, 060401 (2016).

30. Mivehvar, F., Ritsch, H. & Piazza, F. Superradiant topological peierls insulator inside an optical cavity. *Phys. Rev. Lett.* **118**, 073602 (2017).

31. Schlawin, F. & Jaksch, D. Cavity-mediated unconventional pairing in ultracold fermionic atoms. *Phys. Rev. Lett.* **123**, 133601 (2019).

32. Zheng, Z. & Wang, Z. D. Cavity-induced Fulde-Ferrell-Larkin-Ovchinnikov superfluids of ultracold Fermi gases. *Phys. Rev. A* **101**, 023612 (2020).

33. Roux, K., Konishi, H., Helson, V. & Brantut, J.-P. Strongly correlated fermions strongly coupled to light. *Nat. Commun.* **11**, 2974 (2020).

34. Roux, K., Helson, V., Konishi, H. & Brantut, J.-P. Cavity-assisted preparation and detection of a unitary fermi gas. *N. J. Phys.* **23**, 043029 (2021).

35. Watanabe, G., Orso, G., Dalfovo, F., Pitaevskii, L. P. & Stringari, S. Equation of state and effective mass of the unitary fermi gas in a one-dimensional periodic potential. *Phys. Rev. A* **78**, 063619 (2008).

36. Navon, N., Nascimbène, S., Chevy, F. & Salomon, C. The equation of state of a low-temperature fermi gas with tunable interactions. *Science* **328**, 729–732 (2010).

37. Horikoshi, M., Koashi, M., Tajima, H., Ohashi, Y. & Kuwata-Gonokami, M. Ground-state thermodynamic quantities of homogeneous spin-1/2 fermions from the bcs region to the unitarity limit. *Phys. Rev. X* **7**, 041004 (2017).

38. Mottl, R. et al. Roton-type mode softening in a quantum gas with cavity-mediated long-range interactions. *Science* **336**, 1570–1573 (2012).

39. Baumann, K., Mottl, R., Brennecke, F. & Esslinger, T. Exploring symmetry breaking at the Dicke quantum phase transition. *Phys. Rev. Lett.* **107**, 140402 (2011).

40. Brennecke, F. et al. Real-time observation of fluctuations at the driven-dissipative dicke phase transition. *Proc. Natl Acad. Sci. USA* **110**, 11763–11767 (2013).

41. Léonard, J., Morales, A., Zupancic, P., Donner, T. & Esslinger, T. Monitoring and manipulating Higgs and Goldstone modes in a supersolid quantum gas. *Science* **358**, 1415–1418 (2017).

42. Mihaila, B. Lindhard function of a d-dimensional fermi gas. Preprint at https://arxiv.org/abs/1111.5337 (2011).

43. Patel, P. B. et al. Universal sound diffusion in a strongly interacting fermi gas. *Science* **370**, 1222–1226 (2020).

44. Biss, H. et al. Excitation spectrum and superfluid gap of an ultracold fermi gas. *Phys. Rev. Lett.* **128**, 100401 (2022).

45. Fradkin, E., Kivelson, S. A. & Tranquada, J. M. Colloquium: theory of intertwined orders in high temperature superconductors. *Rev. Mod. Phys.* **87**, 457–482 (2015).

46. Zhu, X., Cao, Y., Zhang, J., Plummer, E. W. & Guo, J. Classification of charge density waves based on their nature. *Proc. Natl Acad. Sci. USA* **112**, 2367–2371 (2015).

47. Schlawin, F., Kennes, D. & Sentef, M. Cavity quantum materials. *Appl. Phys. Rev.* **9**, 011312 (2022).

48. Bloch, J., Cavalleri, A., Galitski, V., Hafezi, M. & Rubio, A. Strongly correlated electron-photon systems. *Phys. Rev. Lett.* **606**, 41–48 (2022).

49. Gao, H., Schlawin, F., Buzzi, M., Cavalleri, A. & Jaksch, D. Photoinduced electron pairing in a driven cavity. *Phys. Rev. Lett.* **125**, 053602 (2020).

50. Konishi, H., Roux, K., Helson, V. & Brantut, J.-P. Universal pair polaritons in a strongly interacting fermi gas. *Nature* **596**, 509–513 (2021).

51. Gadway, B., Pertot, D., Reimann, R., Cohen, M. G. & Schneble, D. Analysis of Kapitza-Dirac diffraction patterns beyond the Raman-Nath regime. *Opt. Express* **17**, 19173–19180 (2009).

52. Helson, V. et al. Optomechanical response of a strongly interacting fermi gas. *Phys. Rev. Res.* **4**, 033199 (2022).

53. Ku, M. J. H., Sommer, A. T., Cheuk, L. W. & Zwierlein, M. W. Revealing the superfluid lambda transition in the universal thermodynamics of a unitary fermi gas. *Science* **335**, 563–567 (2012).

54. Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of ultracold atomic fermi gases. *Rev. Modern Phys.* **80**, 1215–1274 (2008).

55. Atland, A & Simons, B. D. *Condensed Matter Field Theory* 2nd edn (Cambridge University Press, 2010).

56. Hu, H., Taylor, E., Liu, X.-J., Stringari, S. & Griffin, A. Second sound and the density response function in uniform superfluid atomic gases. *N. J. Phys.* **12**, 043040 (2010).

Acknowledgements

We acknowledge discussions with T. Donner and T. Esslinger. We thank G. del Pace and T. Bühler for their assistance in the final stages of the experiment. We acknowledge funding from the European Research Council under the European Union Horizon 2020 Research and Innovation Programme (Grant no. 714309) and the Swiss National Science Foundation (Grant no. 184654). F.M. acknowledges financial support from the Austrian Science Fund (Stand-Alone Project P 35891-N).

Funding

Open access funding provided by EPFL Lausanne.

Author information

Authors and Affiliations

1. Institute of Physics, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

   Victor Helson, Timo Zwettler, Kevin Roux, Hideki Konishi & Jean-Philippe Brantut

2. Center for Quantum Science and Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

   Victor Helson, Timo Zwettler, Kevin Roux, Hideki Konishi & Jean-Philippe Brantut

3. Institut für Theoretische Physik, Universität Innsbruck, Innsbruck, Austria

   Farokh Mivehvar, Elvia Colella & Helmut Ritsch

4. Institute of Science and Technology Austria, Klosterneuburg, Austria

   Kevin Roux

5. Department of Physics, Graduate School of Science, Kyoto University, Kyoto, Japan

   Hideki Konishi

Contributions

V.H., T.Z., K.R. and H.K. performed experiments. V.H. and T.Z. processed the data. F.M., E.C. and H.R. performed calculations. J.-P.B. planned and supervised the experiments.

Corresponding author

Correspondence to Jean-Philippe Brantut.

Ethics declarations

Competing interets

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Miguel Angel Bastarrachea-Magnani, Michael Sentef and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Measurement of heating due to the pump beam.

The vertical line depicts the location of threshold for the self-organizing phase transition and the horizontal dashed one marks the superfluid transition for a homogeneously trapped unitary fermi gas. As the pump power is increased, we observe a smooth increase of the gas temperature showing no dramatic behavior around the self-organization phase transition. Error bars represent a standard deviation.

Extended Data Fig. 2 Measured damping of the oscillatory behavior predicted by Eq. (14), which is accounted for by an additional $e^{-t/\tau}$ factor in the equation.

The signal becomes strongly damped as the critical value for the long-range interaction strength is approached. The data shown is part of the set displayed in Fig. 3 of the main text, here taken at unitarity and for $\Delta_c < 0$. In inset, we display the measured phase offset ϕ_0 which features a uniform distribution. Error bars represent a standard deviation.

Extended Data Fig. 3 Critical long-range interaction strength as a function of short-range interaction strength, normalized with respect to the critical strength at unitarity (black circles), compared with the theoretical prediction based on the compressibility (solid orange line).

The data are identical to that of Fig. 2c. Error bars represent a standard deviation.

Supplementary information

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Helson, V., Zwettler, T., Mivehvar, F. *et al.* Density-wave ordering in a unitary Fermi gas with photon-mediated interactions. *Nature* **618**, 716–720 (2023). https://doi.org/10.1038/s41586-023-06018-3

- Received: 09 December 2022
- Accepted: 27 March 2023
- Published: 24 May 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06018-3

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06018-3

- Article
- Open Access
- Published: 14 June 2023

High-brightness scalable continuous-wave single-mode photonic-crystal laser

- Masahiro Yoshida,
- Shumpei Katsuno,
- Takuya Inoue,
- John Gelleta,
- Koki Izumi,
- Menaka De Zoysa,
- Kenji Ishizaki &
- …
- Susumu Noda

Show authors

Nature **volume 618**, pages 727–732 (2023)

- 5351 Accesses

- 53 Altmetric

- Metrics details

Abstract

Realizing large-scale single-mode, high-power, high-beam-quality semiconductor lasers, which rival (or even replace) bulky gas and solid-state

lasers, is one of the ultimate goals of photonics and laser physics. Conventional high-power semiconductor lasers, however, inevitably suffer from poor beam quality owing to the onset of many-mode oscillation[1,2], and, moreover, the oscillation is destabilized by disruptive thermal effects under continuous-wave (CW) operation[3,4]. Here, we surmount these challenges by developing large-scale photonic-crystal surface-emitting lasers with controlled Hermitian and non-Hermitian couplings inside the photonic crystal and a pre-installed spatial distribution of the lattice constant, which maintains these couplings even under CW conditions. A CW output power exceeding 50 W with purely single-mode oscillation and an exceptionally narrow beam divergence of 0.05° has been achieved for photonic-crystal surface-emitting lasers with a large resonant diameter of 3 mm, corresponding to over 10,000 wavelengths in the material. The brightness, a figure of merit encapsulating both output power and beam quality, reaches 1 GW cm^{-2} sr^{-1}, which rivals those of existing bulky lasers. Our work is an important milestone toward the advent of single-mode 1-kW-class semiconductor lasers, which are expected to replace conventional, bulkier lasers in the near future.

Main

Semiconductor lasers boast various beneficial features that cannot be found in other lasers (for example, gas, solid-state and fibre lasers), such as compactness, high efficiency and high controllability, and they are key devices for various applications in modern society, including telecommunications and optical recording. Realizing semiconductor lasers that also operate in a single mode with high output power and high beam quality remains an ultimate yet elusive goal in photonics and laser physics. Such semiconductor lasers are in demand for many emerging applications, including next-generation laser processing, remote sensing, long-range free-space communications and even light propulsion for spaceflight[5,6,7,8]. Conventional semiconductor lasers are limited by the maximum emission area that can support single-mode operation; namely, widening the emission area to increase the output power leads to the onset of many-mode oscillation, which degrades the beam quality[1,2]. Even worse, under continuous-wave (CW) operation, the beam quality is prone to degrade

further due to a thermally induced refractive index distribution inside the resonator, which is one of the critical factors responsible for unstable oscillation[3,4] (Supplementary Text Section 1 has more details).

The photonic-crystal surface-emitting laser (PCSEL)[9,10,11,12,13,14,15] shows potential to overcome the above limitations of conventional semiconductor lasers. The PCSEL achieves lasing oscillation of a two-dimensional standing wave at a singularity (Γ, M and so on) point in its photonic band structure. By tailoring the design of the unit cell of its photonic crystal, the mutual optical couplings inside the photonic crystal can be tuned to enable single-mode oscillation over a large area. One unit-cell design proposed for this purpose is the double lattice[15], in which one lattice point group is shifted from a second in the x and y directions by approximately one quarter of the wavelength in the material. In this double lattice, the strength of in-plane optical coupling, which can be referred to as Hermitian coupling since it is not accompanied by radiation loss, is weakened by destructive interference of waves diffracted by 180° and 90° at each of the two lattice points. Consequently, optical losses of higher-order modes from the periphery of the resonator increase compared with that of the fundamental mode, resulting in a wider threshold gain margin between these modes and thus, more stable oscillation in the fundamental mode. Based on this concept, CW lasing oscillation with an output power of 7 W and a brightness of 180 MW cm^{-2} sr^{-1} was experimentally demonstrated using PCSELs with circular resonant diameters of 800 μm (ref. [15]).

Following the above developments, a design guideline to realize single-mode oscillation over areas of even larger (≥3 mm) diameters was recently reported[16] based on the control of not only the Hermitian coupling described above, but also non-Hermitian coupling, which accompanies radiation loss. In addition, another related approach toward realizing scalable single-mode photonic-crystal lasers was also reported[17], although the resonant (emission) size achieved in experiments therein was less than approximately 64 μm. (Supplementary Text Section 2 has a brief comparison between these two approaches).

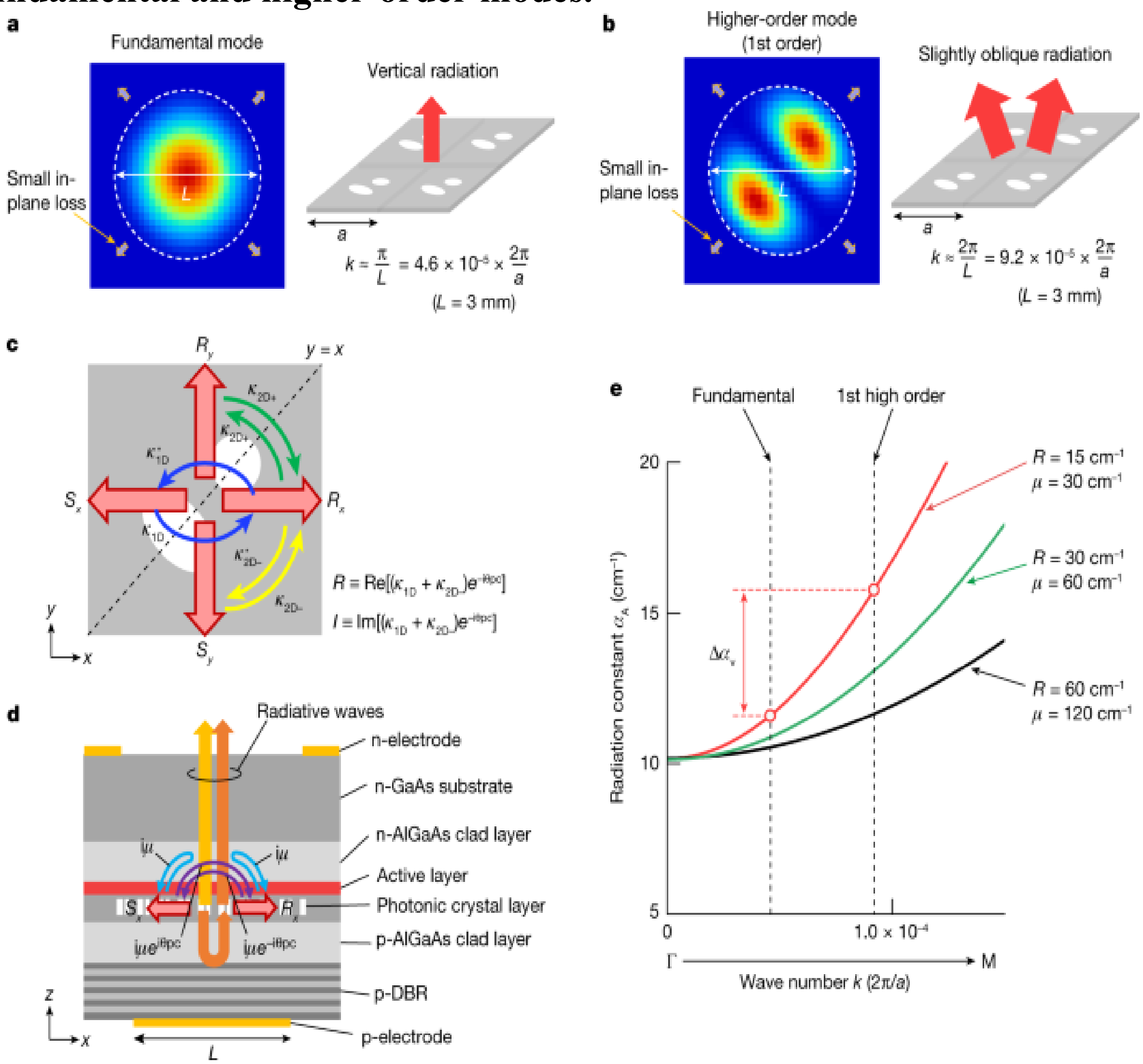
The most important outstanding challenges are, therefore, twofold. One is to investigate whether single-mode operation of PCSELs can be truly scaled to

extremely large (≥3 mm) diameters; the other is to investigate whether single-mode operation can be maintained even under CW conditions, where disruptive thermal effects appear. Here, we first show that it is indeed possible to realize single-mode oscillation even in a PCSEL with a 3-mm diameter, corresponding to over 10,000 wavelengths in the material, by simultaneously controlling the Hermitian and non-Hermitian couplings. Then, we introduce a lattice-constant distribution to compensate for the thermal effects and thereby maintain the controlled Hermitian and non-Hermitian couplings even under CW conditions. By doing so, we finally experimentally achieve 50-W CW operation in a single mode (with a single wavelength) with a very narrow beam divergence angle of 0.05° (corresponding to a beam quality $M^2 \approx 2.36$). The brightness of the developed laser reaches 1 GW cm^{-2} sr^{-1}, which is more than one order of magnitude greater than that of conventional semiconductor lasers and even rivals those of existing bulky gas and solid-state lasers. The strategies demonstrated here are expected to be applicable to scaling up the diameter of the device to 10 mm, leading to the 1-kW class, high-beam-quality operation of PCSELs.

Hermitian and non-Hermitian-controlled PCSEL

First, we describe the strategy to realize single-mode oscillation in a large-scale PCSEL. The left panels of Fig. 1a,b show typical electric-field distributions of the fundamental and higher-order modes in a PCSEL with a diameter of *L*. As *L* increases, the in-plane losses (that is, light escaping from the periphery of the resonator) of both the fundamental and higher-order modes converge toward zero, and thus, the ability to discriminate between these modes via in-plane loss greatly diminishes. Accordingly, we instead consider mode discrimination via vertical radiation loss (that is, the radiation constant), which remains high even when *L* becomes large. As illustrated in the right panels of Fig. 1a,b, the (first) higher-order mode is double lobed, and consequently, it has a slightly larger in-plane wave number than the single-lobed fundamental mode. Our strategy is to make the radiation constant at the wave number corresponding to the (first) higher-order mode sufficiently larger than that of the fundamental mode by controlling the Hermitian and non-Hermitian couplings in a double-lattice photonic crystal.

Fig. 1: Control of Hermitian and non-Hermitian optical couplings inside the photonic crystal to increase the threshold gain margin between the fundamental and higher-order modes.



a, Typical electric-field intensity distribution (left panel) and schematic of radiation (right panel) of the fundamental mode in a PCSEL with a diameter of L. **b**, Typical electric-field intensity distribution (left panel) and schematic of radiation (right panel) of a higher-order mode in the same PCSEL. The higher-order mode emits a multilobed beam in a slightly oblique direction. **c**, Hermitian couplings between the four fundamental waves (R_x, S_x, R_y and S_y) inside a PCSEL with a double-lattice photonic crystal. **d**, Schematic of the cross-sectional structure of the PCSEL and the non-Hermitian couplings between the fundamental waves via radiative waves inside the device, where

the couplings between R_x and S_x are illustrated as an example. A backside reflector (p-type distributed Bragg reflector (p-DBR)) is used to control the magnitude of the non-Hermitian coupling coefficient. **e**, Calculated radiation constants of mode A as functions of k when the values of R and μ are changed.

Figure 1c,d illustrates the Hermitian and non-Hermitian couplings of four fundamental waves, labelled R_x, S_x, R_y and S_y, in a double-lattice structure, where the former and latter couplings do not and do accompany radiation loss, respectively. In Fig. 1c, the coefficients κ_{1D} and $\kappa_{2D\pm}$ express the strengths of Hermitian coupling at angles of 180° and ±90°, respectively. The self-coupling of the four fundamental waves without radiation loss is expressed by κ_{11} (not shown in the figure). In Fig. 1d, the coefficients iμ and i$\mu e^{\pm i\theta_{pc}}$ express the strengths of non-Hermitian couplings at angles of 0° (that is, self-coupling) and 180°, respectively, via the radiative waves, where μ is the magnitude of the non-Hermitian coupling coefficient, θ_{pc} is a phase change associated with ±180° non-Hermitian coupling and i is the imaginary unit.

Owing to such optical couplings among the four fundamental waves, four resonant cavity modes are constructed. In a double-lattice photonic crystal with mirror symmetry about the axis of $y = x$, two of these modes (labelled A and C) are antisymmetric modes whose electric-field vectors are antisymmetric about this axis, and the remaining two modes (labelled B and D) are symmetric modes whose electric-field vectors are symmetric about this axis. The radiation constant, as well as the frequency, of each mode is analytically derived as a function of the in-plane wave number k using three-dimensional coupled-wave theory[16,18]. The frequency and radiation constant (δ_A, α_A) of mode A, which has the lowest radiation constant among all four modes, and those (δ_C, α_C) of its counterpart mode C are expressed as (Methods has the details of the derivation)

$${\delta }_{{\rm{A}},{\rm{C}}}+{\rm{i}}\frac{{\alpha }_{{\rm{A}},{\rm{C}}}}{2}={\kappa }_{11}+{\kappa }_{2{\rm{D}}+}+{\rm{i}}\mu \mp \sqrt{[{R}+{\rm{i}}{I}+{\rm{i}}\mu][{R}-{\rm{i}}{I}+{\rm{i}}\mu]+{\left(\frac{{k}}{\sqrt{2}}\right)}^{2}}.$$

(1)

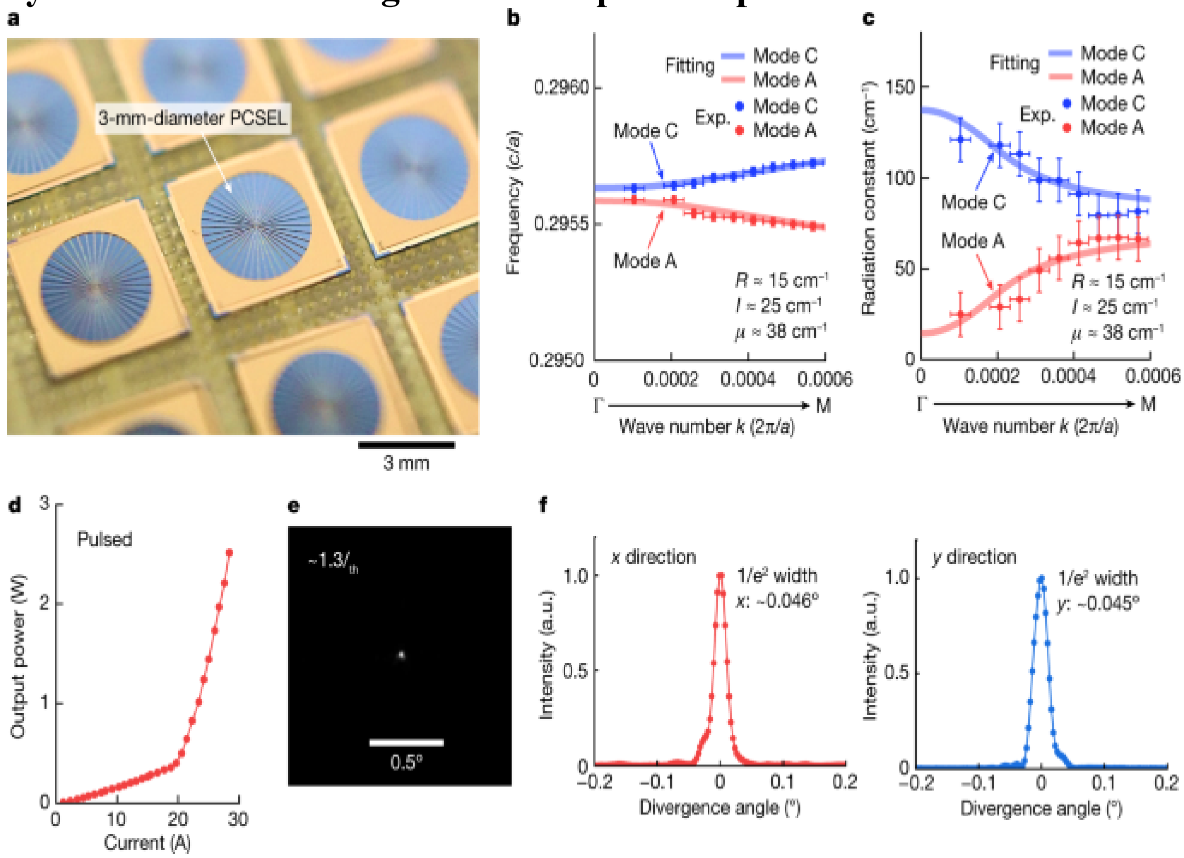
Here, $\delta_A + i\alpha_A/2$ and $\delta_C + i\alpha_C/2$ correspond to the negative and positive square-root terms, respectively, and the in-plane wave number k is taken along the Γ–M direction (that is, ${k}_{x}={k}_{y}=k/\sqrt{2}$), so that the symmetry of the electric fields coincides with the symmetry of the photonic crystal along the line of $y = x$. In equation (1), R and I correspond to the real and imaginary parts of the Hermitian coupling coefficient, respectively; namely, $R\equiv {\rm{Re}}\left[\left({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}\right){{\rm{e}}}^{{\rm{-i}}{\theta }_{{\rm{pc}}}}\right]$ and $I\equiv {\rm{Im}}\left[\left({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}\right){{\rm{e}}}^{{\rm{-i}}{\theta }_{{\rm{pc}}}}\right]$. R expresses the overall strength of in-plane feedback of combined 180° and 90° diffractions at the Γ point, which determines the size of the frequency gap between modes A and C. I determines the phase of the in-plane electric fields, specifically the position of the electric-field node with respect to the position of the air holes, and consequently, it determines the degree of cancellation of the vertical radiation in mode A at the Γ point. (Details on R and I are explained in Supplementary Text Section 3).

Figure 1e shows α_A as a function of k calculated using equation (1) for several selections of R and μ, while I was adjusted so that α_A at $k = 0$ was identical in all cases. Evidently, decreasing R and μ simultaneously while maintaining their balance increases the curvature of dispersion around the Γ point, resulting in an abrupt change of α_A with respect to k. Consequently, the threshold margin $\Delta\alpha_v$ between the fundamental and higher-order modes near the Γ point can be increased, whereupon single-mode oscillation is expected even for large 3-mm-diameter PCSELs (Methods has a note on $\Delta\alpha_v$).

Based on the above strategy, we developed a 3-mm-diameter PCSEL, whose R and I values were controlled by changing the lattice separation (d) and the balance of air-hole sizes (x) of a double-lattice structure (Fig. 1c and Supplementary Text Section 4) and whose μ value was controlled by changing the thickness of the p-AlGaAs clad layer, which affects the degree of optical interference between front side-emitted and back side-reflected

radiative waves[16] (Fig. 1d and Supplementary Text Section 4). Figure 2a shows finished 3-mm-diameter PCSELs fabricated based on an air hole-retained metal–organic vapour-phase epitaxy regrowth technique[19]. A mesh-window n-electrode was deposited onto an n-GaAs substrate (emission side) for uniform current injection across the entire 3-mm-diameter area.

Fig. 2: Development of a 3-mm-diameter PCSEL and demonstration of very narrow beam divergence under pulsed operation.



a, Photograph of the fabricated 3-mm-diameter PCSEL chips. **b**,**c**, Frequencies (**b**) and radiation constants (**c**) as functions of k for modes A and C. The dots indicate experimental data (Exp.) obtained from the peak wavelengths and line widths of the spontaneous emission spectrum at each wave number. The solid lines are theoretical curves based on equation (1), where the best-fit results are $R = 15\ \mathrm{cm}^{-1}$, $I = 25\ \mathrm{cm}^{-1}$ and $\mu = 38\ \mathrm{cm}^{-1}$. The vertical error bars in **b** and **c** represent measurement uncertainty associated with the central frequency and the spectral line width, and the horizontal error bars represent measurement uncertainty associated with the wave number. **d**, Measured I–L characteristics under pulsed conditions. The pulse

width and repetition frequency were set to 200 ns and 200 Hz, respectively. **e**, Measured FFP at an injection current of $1.3I_{th}$. **f**, Cross-sectional intensity profiles of the FFP along the *x* and *y* directions. A very narrow $1/e^2$ divergence angle of 0.045° was obtained owing to the achievement of small yet appropriately balanced values of *R* and *μ* as described in the text. a.u., arbitrary unit.

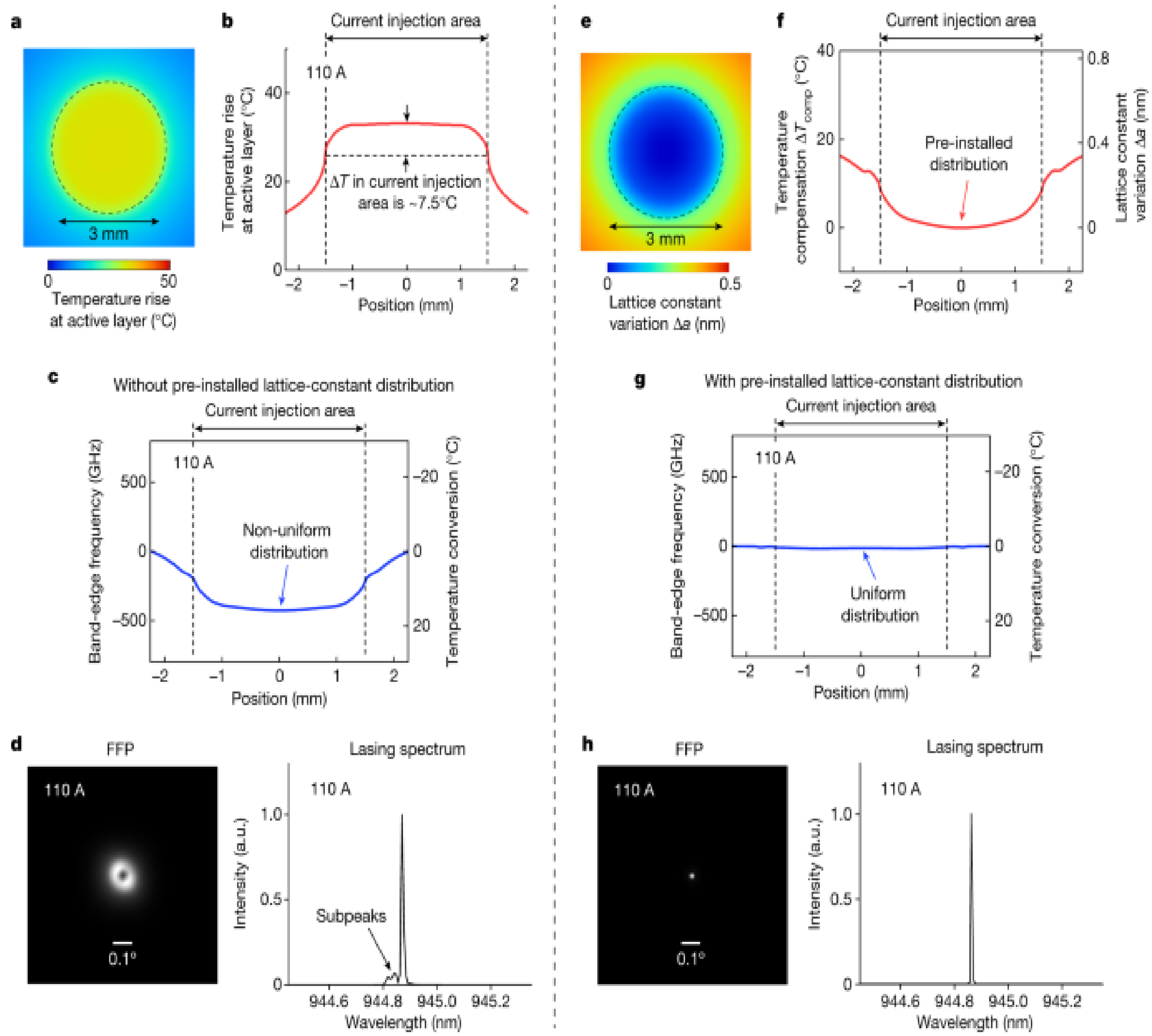
We first measured the frequencies and radiation constants of modes A and C around the Γ point of the fabricated device as plotted in Fig. 2b,c and then, estimated the coupling coefficients *R*, *I* and *μ* by fitting the analytical values given by equation (1) to their experimental values (Methods has details). The best-fit results were $R \approx 15\ \text{cm}^{-1}$, $I \approx 25\ \text{cm}^{-1}$ and $\mu \approx 38\ \text{cm}^{-1}$; these values correspond to the case of the red line in Fig. 1e, in which a large threshold margin $\Delta\alpha_v$ is expected. Then, we measured lasing characteristics of the PCSEL under pulsed conditions. Here, the PCSEL was not mounted to a heat sink, which limited the maximum tolerable injection current and hence output power even under pulsed conditions. The current–light output (*I*–*L*) characteristics measured at room temperature in Fig. 2d show that lasing oscillation occurred at a threshold injection current of $I_{th} \approx 20$ A. Figure 2e shows a far-field pattern (FFP) of the emitted beam above the threshold ($1.3\ I_{th}$). As indicated by the cross-sectional profile of the FFP (Fig. 2f), a very narrow beam divergence angle of 0.045° was achieved, which we attribute to the small yet appropriately balanced values of *R* and *μ* as described above.

High-brightness CW single-mode PCSEL

CW current injection induces a spatially non-uniform temperature distribution inside the PCSEL due to the accumulation of heat. We simulated the effects of heat accumulation on the lasing characteristics of a 3-mm-diameter PCSEL under CW conditions based on a self-consistent analysis[20] of the interaction among photons, carriers and thermal effects (Supplementary Text Section 5 for details). Figure 3a,b shows the calculated in-plane temperature distribution near the active layer at a sufficiently large CW injection current of 110 A (chosen upon consideration of the experimental conditions described later). As shown, the temperature at the

centre of the current injection area becomes higher than at the periphery, which results in the downward convex-shaped band-edge frequency distribution shown in Fig. 3c via a change in refractive index. This frequency distribution perturbs the electric field of the fundamental mode, and consequently, it induces multimodal behaviour and broadens the emitted beam, as shown in Fig. 3d. To suppress such unwanted effects, we introduce a spatial variation $\Delta a(x,y)$ to the lattice constant of the photonic crystal, which compensates for a temperature distribution $\Delta T_{\mathrm{comp}}(x,y)$, as shown in Fig. 3e,f (Methods has details). Figure 3g shows the band-edge frequency distribution calculated at an injection current of 110 A following the introduction of this lattice-constant distribution. This figure clearly shows that a uniform frequency distribution is obtained, owing to the mutual cancellation of the temperature-induced and lattice constant-induced changes of the band-edge frequency. As a result, the emission of a single-mode beam with a very narrow divergence angle is expected to be obtained, as shown in Fig. 3h.

Fig. 3: Introduction of a lattice-constant distribution to realize single-mode oscillation of a 3-mm-diameter PCSEL even under CW operation.



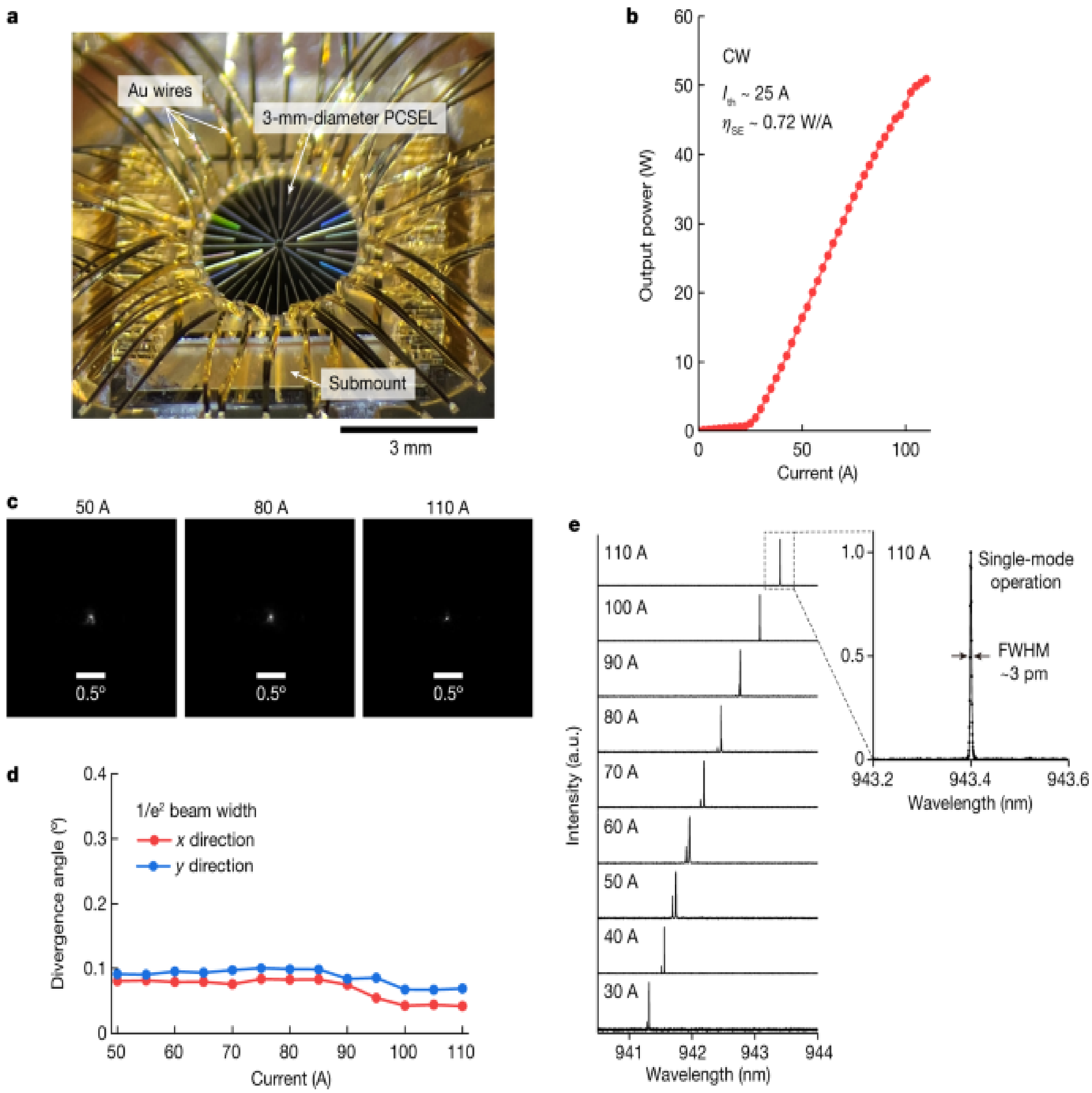
a,**b**, Calculated in-plane distribution (**a**) and its profile (**b**) of temperature rise near the active layer at a CW injection current of 110 A. From these calculations, the temperature difference Δ*T* between the centre and the edge of the current injection area was estimated to be 7.5 °C. **c**, Band-edge frequency distribution of a 3-mm-diameter PCSEL without a pre-installed lattice-constant distribution. The right axis shows the temperature change converted from the frequency. A non-uniform band-edge frequency distribution is formed due to the temperature-induced refractive index distribution. **d**, FFP and lasing spectrum of a 3-mm-diameter PCSEL without a pre-installed lattice-constant distribution. **e**, In-plane distribution of the lattice-constant variation to compensate for the temperature distribution that appears at an injection current of 110 A. **f**, Lattice-constant variation and corresponding temperature compensation as functions of

position. **g**, Band-edge frequency distribution of a 3-mm-diameter PCSEL with the pre-installed lattice-constant distribution. The right axis shows the temperature change converted from the frequency. A uniform frequency distribution is realized by the mutual cancellation of the temperature-induced and lattice constant-induced changes to the band-edge frequency. **h**, FFP and lasing spectrum of a 3-mm-diameter PCSEL with the pre-installed lattice-constant distribution.

Applying the above strategy, we developed the 3-mm-diameter PCSEL with the pre-installed lattice-constant distribution. Figure 4a shows a finished 3-mm-diameter PCSEL mounted on a package. The coupling coefficients of the device were estimated as $R \approx 24\ \mathrm{cm}^{-1}$, $I \approx 14\ \mathrm{cm}^{-1}$ and $\mu \approx 44\ \mathrm{cm}^{-1}$ (Methods has details). Figure 4b shows *I*–*L* characteristics of the 3-mm-diameter PCSEL under CW conditions. The threshold current was 25 A, and the slope efficiency was approximately $0.72\ \mathrm{W\,A}^{-1}$. A CW output power exceeding 50 W was obtained from the single-chip PCSEL at injection currents of 100–110 A. Figure 4c shows the FFPs at several injection currents, and Fig. 4d shows the $1/\mathrm{e}^2$ beamwidths of FFPs as functions of the injection current. Remarkably, the divergence angles in the *x* and *y* directions became minimal (0.05°) at 100–110 A, where the frequency distributions due to the temperature-induced refractive index change and the pre-installed lattice-constant distribution were designed to cancel each other out. We note that the divergence angle was slightly larger in the *y* direction than in the *x* direction due to a small residual side lobe with an intensity of approximately $1/\mathrm{e}^2$ of that of the main peak; this side lobe can be eliminated in the future by further optimizing the pre-installed lattice-constant distribution. From the divergence angles, including the small residual side lobe, the beam quality M^2 was estimated to be 2.36. The CW laser brightness, evaluated using the measured output power and FFP widths at 110 A, was $1\ \mathrm{GW\,cm}^{-2}\,\mathrm{sr}^{-1}$. Furthermore, the laser spectra at several injection currents are shown in Fig. 4e. Although several modes were seen to oscillate at lower injection currents, single-mode oscillation was achieved at injection currents of 100–110 A (corresponding to a CW output power of around 50 W). The measured spectral width at this injection current was 3 pm, which was limited by the spectral resolution of our spectrometer. As this resolution was finer than the predicted spectral spacing between the fundamental and next higher-order modes, we may say that purely single-mode oscillation was achieved. We

note that $M^2 \geq 2$ in spite of single-mode oscillation is due to the super-Gaussian electromagnetic-field intensity profile caused predominantly by the uniform current injection. The dependence of the laser spectrum on the injection current agrees with that of the beam divergence angle plotted in Fig. 4d. These results show that the pre-installed lattice-constant distribution together with the control of Hermitian and non-Hermitian couplings inside the double-lattice structure has contributed to the realization of purely single-mode, high beam quality, high-power CW operation of an ultra-large-area PCSEL.

Fig. 4: Demonstration of 50-W single-mode CW operation of a 3-mm-diameter PCSEL.



a, Photograph of the 3-mm-diameter PCSEL, which was mounted on a package via a submount with high thermal conductivity in the p-side-down configuration. To inject a high CW current, many Au wires were bonded to the n-side electrode of the PCSEL and to the submount surface (p side). **b**, *I*–*L* characteristics under CW conditions, where the heat sink was cooled by water at a temperature of 6 °C during measurements. **c**, FFPs at injection currents of 50, 80 and 110 A. **d**, The $1/e^2$ beam divergence angle versus injection current along the *x* and *y* directions. Very narrow beam divergence angles of 0.05° were obtained. **e**, Lasing spectra measured at various injection currents. A narrow spectral width was observed. Single-mode oscillation was realized at injection currents of 100–110 A. FWHM, full width at half maximum.

Conclusions

We have developed large-scale PCSELs by control of Hermitian and non-Hermitian couplings to suppress the oscillation of higher-order modes, and we have introduced a lattice-constant distribution to maintain these controlled couplings even under CW operation. By doing so, we have realized 50-W single-mode (or single-wavelength) oscillation of a PCSEL with an ultra-large diameter of 3 mm, corresponding to over 10,000 wavelengths in the material. The 50-W CW output power and a very narrow beam divergence of 0.05° ($M^2 \approx 2.36$) correspond to a brightness of 1 GW cm^{-2} sr^{-1}, which rivals those of existing bulky lasers. Controlling the Hermitian and non-Hermitian coupling coefficients (R, I and μ) and introducing a lattice-constant distribution suitable for devices of even larger scales (for example, 10-mm diameters) are expected to contribute to the realization of 1-kW class, high-beam-quality operation of PCSELs. Our work is an important milestone toward the replacement of conventional, bulkier solutions and toward innovation in a wide variety of industrial applications, from smart material processing to aerospace applications.

Methods

Derivation of frequencies and radiation constants of modes A and C

The frequency and radiation constant (δ_A, α_A) of mode A, which has the lowest radiation constant among all four modes, and those (δ_C, α_C) of its counterpart mode C can be obtained by solving the following equation based on three-dimensional coupled-wave theory[16,18]:

$$\begin{array}{l}\left(\delta +{\rm{i}}\frac{\alpha }{2}\right)\,\left(\begin{array}{c}{R}_{x}\\ {S}_{x}\\ {R}_{y}\\ {S}_{y}\end{array}\right)\,=\,\left[\left(\begin{array}{cccc}{\kappa }_{11} & {\kappa }_{1{\rm{D}}} & {\kappa }_{2{\rm{D}}+} & {\kappa }_{2{\rm{D}}-}\\ {\kappa }_{1{\rm{D}}}^{\ast } & {\kappa }_{11} & {\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{2{\rm{D}}+}\\ {\kappa }_{2{\rm{D}}+} & {\kappa }_{2{\rm{D}}-} & {\kappa }_{11} & {\kappa }_{1{\rm{D}}}\\ {\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{2{\rm{D}}+} & {\kappa }_{1{\rm{D}}}^{\ast } & {\kappa }_{11}\end{array}\right)\right.\\ \,\,\,+\,\left(\begin{array}{cccc}{\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & 0 & 0\\ {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu & 0 & 0\\ 0 & 0 & {\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}\\ 0 & 0 & {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu \end{array}\right)\\ \,\,\,\left.+\,\left(\begin{array}{cccc}{k}_{x} & 0 & 0 & 0\\ 0 & -{k}_{x} & 0 & 0\\ 0 & 0 & {k}_{y} & 0\\ 0 & 0 & 0 & -{k}_{y}\end{array}\right)\right]\,\left(\begin{array}{c}{R}_{x}\\ {S}_{x}\\ {R}_{y}\\ {S}_{y}\end{array}\right).\end{array}$$
(2)

The first and second terms on the right-hand side of equation (2) represent the Hermitian and non-Hermitian coupling processes described in Fig. 1c,d, respectively. The third term on the right-hand side of equation (2) represents the deviation of the wave number from the Γ point in an arbitrary direction represented by wave numbers k_x and k_y, which induces a change in the eigenfrequency of each mode.

Here, we consider the eigenfrequencies of the modes in the Γ–M direction ($k_x = k_y = k/\sqrt{2}$), which is parallel to the axis of symmetry of the double-lattice photonic crystal ($y = x$). Based on the symmetry along y

$= x$, the coupled-wave matrices on the right-hand side of equation (2) can be block diagonalized using the basis-transformation matrix P:

$$P=\frac{1}{\sqrt{2}}\left(\begin{array}{cccc}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\end{array}\right),$$
(3)

as

$$\begin{array}{c}{P}^{-1}CP\,=\,\left(\begin{array}{cccc}{\kappa }_{11}+{\kappa }_{2{\rm{D}}+} & {\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-} & 0 & 0\\ {\kappa }_{1{\rm{D}}}^{\ast }+{\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{11}+{\kappa }_{2{\rm{D}}+} & 0 & 0\\ 0 & 0 & {\kappa }_{11}-{\kappa }_{2{\rm{D}}+} & {\kappa }_{1{\rm{D}}}-{\kappa }_{2{\rm{D}}-}\\ 0 & 0 & {\kappa }_{1{\rm{D}}}^{\ast }-{\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{11}-{\kappa }_{2{\rm{D}}+}\end{array}\right)\\ \,\,+\,\left(\begin{array}{cccc}{\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & 0 & 0\\ {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu & 0 & 0\\ 0 & 0 & {\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}\\ 0 & 0 & {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu \end{array}\right)\\ \,\,+\,\frac{1}{\sqrt{2}}\left(\begin{array}{cccc}k & 0 & 0 & 0\\ 0 & -k & 0 & 0\\ 0 & 0 & k & 0\\ 0 & 0 & 0 & -k\end{array}\right).\end{array}$$
(4)

Then, the coupled-wave equation (2) can be divided into the following two forms:

$$\begin{array}{l}\left(\delta +{\rm{i}}\frac{\alpha }{2}\right)\,\left(\begin{array}{c}{R}_{x}+{R}_{y}\\ {S}_{x}+{S}_{y}\end{array}\right)=\left[\left(\begin{array}{cc}{\kappa }_{11}+{\kappa }_{2{\rm{D}}+} & {\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}\\ {\kappa }_{1{\rm{D}}}^{\ast }+{\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{11}+{\kappa }_{2{\rm{D}}+}\end{array}\right)\right.+\left(\begin{array}{cc}

{\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}\\ {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu \end{array}\right)\\ \,\,\,\,\left.+\,\frac{1}{\sqrt{2}}\left(\begin{array}{cc}k & 0\\ 0 & -k\end{array}\right)\right]\,\left(\begin{array}{c}{R}_{x}+{R}_{y}\\ {S}_{x}+{S}_{y}\end{array}\right)\end{array}$$
(5)

and

$$\begin{array}{l}\left(\delta +{\rm{i}}\frac{\alpha }{2}\right)\,\left(\begin{array}{c}{R}_{x}-{R}_{y}\\ {S}_{x}-{S}_{y}\end{array}\right)=\left[\left(\begin{array}{cc}{\kappa }_{11}-{\kappa }_{2{\rm{D}}+} & {\kappa }_{1{\rm{D}}}-{\kappa }_{2{\rm{D}}-}\\ {\kappa }_{1{\rm{D}}}^{\ast }-{\kappa }_{2{\rm{D}}-}^{\ast } & {\kappa }_{11}-{\kappa }_{2{\rm{D}}+}\end{array}\right)+\left(\begin{array}{cc}{\rm{i}}\mu & {\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}\\ {\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}} & {\rm{i}}\mu \end{array}\right)\right.\\ \,\,\,\,\left.+\,\frac{1}{\sqrt{2}}\left(\begin{array}{cc}k & 0\\ 0 & -k\end{array}\right)\right]\,\left(\begin{array}{c}{R}_{x}-{R}_{y}\\ {S}_{x}-{S}_{y}\end{array}\right),\end{array}$$
(6)

where equation (5) gives the coupling between a pair of electric-field vectors $R_x + R_y$ and $S_x + S_y$, which leads to the formation of modes A and C (Supplementary Fig. 1).

The frequencies and radiation constants of modes A and C can be then derived from equation (5) as follows:

$${\delta }_{{\rm{A}},{\rm{C}}}+{\rm{i}}\frac{{\alpha }_{{\rm{A}},{\rm{C}}}}{2}={\kappa }_{11}+{\kappa }_{2{\rm{D}}+}+{\rm{i}}\mu \mp \sqrt{[({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-})+{\rm{i}}\mu {{\rm{e}}}^{{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}][{({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-})}^{\ast }+{\rm{i}}\mu {{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}]+

{\left(\frac{{k}}{\sqrt{2}}\right)}^{2}}={\kappa }_{11}+{\kappa }_{2{\rm{D}}+}+{\rm{i}}\mu \mp \sqrt{[({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}){{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}}+{\rm{i}}\mu][{\{({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}){{\rm{e}}}^{-{\rm{i}}{\theta }_{{\rm{p}}{\rm{c}}}}\}}^{\ast }+{\rm{i}}\mu]+{\left(\frac{{k}}{\sqrt{2}}\right)}^{2}}={\kappa }_{11}+{\kappa }_{2{\rm{D}}+}+{\rm{i}}\mu \mp \sqrt{[{R}+{\rm{i}}{I}+{\rm{i}}\mu][{R}-{\rm{i}}{I}+{\rm{i}}\mu]+{\left(\frac{{k}}{\sqrt{2}}\right)}^{2}},$$
(7)

where $\delta_A + i\alpha_A/2$ of mode A and $\delta_C + i\alpha_C/2$ of mode C correspond to the negative and positive square-root terms, respectively, and *R* and *I* are defined as \(R\equiv {\rm{Re}}\left[\left({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}\right){{\rm{e}}}^{{\rm{-i}}{\theta }_{{\rm{pc}}}}\right]\) and \(I\equiv {\rm{Im}}\left[\left({\kappa }_{1{\rm{D}}}+{\kappa }_{2{\rm{D}}-}\right){{\rm{e}}}^{{\rm{-i}}{\theta }_{{\rm{pc}}}}\right]\), respectively.

Note on the threshold gain margin $\Delta\alpha_v$

It is difficult to specify a general value of $\Delta\alpha_v$ sufficient for single-mode oscillation in PCSELs. However, we have found that increasing $\Delta\alpha_v$ by simultaneously reducing *R* and *μ* contributes to the preservation of single-mode oscillation even in the presence of a non-uniform in-plane refractive index distribution borne by various physical phenomena. These findings will be reported separately.

Estimation of the coupling coefficients *R*, *I* and *μ* of the fabricated devices

To estimate the coupling coefficients *R*, *I* and *μ* of the fabricated device shown in Fig. 2a, we derived the photonic band structure around the Γ point by measuring the subthreshold spontaneous emission spectra at various radiation angles (corresponding to in-plane wave numbers), whose peak emission wavelengths and line widths corresponded to the band frequencies

and radiation constants, respectively. The frequencies and radiation constants of modes A and C are plotted in Fig. 2b,c. R, I and μ were then estimated by fitting the analytical frequencies and radiation constants given by equation (1) to their measured values.

On the other hand, it was difficult to estimate the coupling coefficients of the fabricated device shown in Fig. 4a by directly measuring the photonic band structure around the Γ point due to the pre-installed lattice-constant distribution. Thus, the coupling coefficients were instead estimated by evaluating the shape of the embedded air holes.

Design of the pre-installed lattice-constant distribution

The distribution of the lattice-constant variation $\Delta a(x,y)$, which compensates for a temperature distribution $\Delta T_{\rm comp}(x,y)$, is determined as follows:

$$\frac{\Delta a(x,y)}{a}\,=\,\frac{\Delta {T}_{{\rm{comp}}}(x,y)}{{n}_{{\rm{eff}}}}\frac{{\rm{d}}n}{{\rm{d}}T}.$$
(8)

Here, a is the original lattice constant, $n_{\rm eff}$ is the effective refractive index of the photonic crystal at room temperature, and $\mathrm{d}n/\mathrm{d}T$ is the rate of change of refractive index with respect to temperature. Based on equation (8) and the self-consistent analysis of photon–carrier–thermal interactions (Supplementary Text Section 5), we introduced the lattice-constant distribution shown in Fig. 3e,f.

Data availability

The data that support the plots within this paper and other findings of this study are available within this article and Supplementary Information, and they are also available from the corresponding author upon request.

Code availability

The details of three-dimensional coupled-wave theory simulations are available in Supplementary Information, and their associated codes are available from the corresponding author upon reasonable request.

References

1. Crump, P. et al. Experimental and theoretical analysis of the dominant lateral waveguiding mechanism in 975 nm high power broad area diode lasers. *Semicond. Sci. Technol.* **27**, 045001 (2012).

2. Chang-Hasnain, C. J. et al. Transverse mode characteristics of vertical cavity surface-emitting lasers. *Appl. Phys. Lett.* **57**, 218–220 (1990).

3. Winterfeldt, M., Crump, P., Wenzel, H., Erbert, G. & Tränkle, G. Experimental investigation of factors limiting slow axis beam quality in 9xx nm high power broad area diode lasers. *J. Appl. Phys.* **116**, 063103 (2014).

4. Rauch, S. et al. Impact of longitudinal refractive index change on the near-field width of high-power broad-area diode lasers. *Appl. Phys. Lett.* **110**, 263504 (2017).

5. Bäuerle, D. *Laser Processing and Chemistry*. (Springer Science & Business Media, 2013).

6. Hecht, J. Lidar for self-driving cars. *Opt. Photon. News* **29**, 26–33 (2018).

7. Kaushal, H. & Kaddoum, G. Optical communication in space: challenges and mitigation techniques. *IEEE Commun. Surv. Tutor.* **19**, 57–96 (2017).

8. Starshot. *Breakthrough Initiatives* https://breakthroughinitiatives.org/Initiative/3 (2018).

9. Imada, M. et al. Coherent two-dimensional lasing action in surface-emitting laser with triangular-lattice photonic crystal structure. *Appl. Phys. Lett.* **75**, 316–318 (1999).

10. Riechel, S. et al. A nearly diffraction limited surface emitting conjugated polymer laser utilizing a two-dimensional photonic band structure. *Appl. Phys. Lett.* **77**, 2310–2312 (2000).

11. Noda, S., Yokoyama, M., Imada, M., Chutinan, A. & Mochizuki, M. Polarization mode control of two-dimensional photonic crystal laser by unit cell structure design. *Science* **293**, 1123–1125 (2001).

12. Kim, M. et al. Surface-emitting photonic-crystal distributed-feedback laser for the midinfrared. *Appl. Phys. Lett.* **88**, 191105 (2006).

13. Hirose, K. et al. Watt-class high-power, high-beam-quality photonic-crystal lasers. *Nat. Photonics* **8**, 406–411 (2014).

14. Noda, S., Kitamura, K., Okino, T., Yasuda, D. & Tanaka, Y. Photonic-crystal surface-emitting lasers: review and introduction of modulated-photonic crystals. *IEEE J. Sel. Top. Quantum Electron.* **23**, 4900107 (2017).

15. Yoshida, M. et al. Double-lattice photonic-crystal resonators enabling high-brightness semiconductor lasers with symmetric narrow-divergence beams. *Nat. Mater.* **18**, 121–128 (2019).

16. Inoue, T. et al. General recipe to realize photonic-crystal surface-emitting lasers with 100-W-1-kW single-mode operation. *Nat. Commun.* **13**, 3262 (2022).

17. Contractor, R. et al. Scalable single-mode surface-emitting laser via open-Dirac singularities. *Nature* **608**, 692–698 (2022).

18. Liang, Y., Peng, C., Sakata, K., Iwahashi, S. & Noda, S. Three-dimensional coupled-wave model for square-lattice photonic crystal lasers with transverse electric polarization: a general approach. *Phys. Rev. B* **84**, 195119 (2011).

19. Yoshida, M. et al. Fabrication of photonic crystal structures by tertiary-butyl arsine-based metal–organic vapor-phase epitaxy for photonic crystal lasers. *Appl. Phys. Express* **9**, 062702 (2016).

20. Katsuno, S. et al. Self-consistent analysis of photonic-crystal surface-emitting lasers under continuous-wave operation. *Opt. Express* **29**, 25118–25132 (2021).

Acknowledgements

This work was partially supported by the project of the Council for Science, Technology and Innovation; the Cross Ministerial Strategic Innovation Promotion Program; Photonics and Quantum Technology for Society 5.0 (finding agency: QST); and the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research 22H04915).

Author information

Author notes

1. These authors contributed equally: Masahiro Yoshida, Shumpei Katsuno, Takuya Inoue

Authors and Affiliations

1. Department of Electronic Science and Engineering, Kyoto University, Kyoto, Japan

   Masahiro Yoshida, Shumpei Katsuno, John Gelleta, Koki Izumi & Susumu Noda

2. Photonics and Electronics Science and Engineering Center, Kyoto University, Kyoto, Japan

   Takuya Inoue, Menaka De Zoysa, Kenji Ishizaki & Susumu Noda

Contributions

S.N. supervised the entire project. M.Y. and S.K. performed the device fabrications and the measurements with K.Izumi, M.D.Z. and K.Ishizaki. T.I. performed the theoretical analysis with M.Y. and J.G. S.K., T.I. and

M.Y. performed the numerical simulation. S.N. and M.Y. discussed the results and wrote the paper with T.I. and J.G.

Corresponding author

Correspondence to Susumu Noda.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information Supplementary Figs. 1–3, Table 1 and Text.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Yoshida, M., Katsuno, S., Inoue, T. *et al.* High-brightness scalable continuous-wave single-mode photonic-crystal laser. *Nature* **618**, 727–732 (2023). https://doi.org/10.1038/s41586-023-06059-8

- Received: 12 January 2023
- Accepted: 05 April 2023
- Published: 14 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06059-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06059-8

- Article
- Open Access
- Published: 21 June 2023

Entropic repulsion of cholesterol-containing layers counteracts bioadhesion

- Jens Friedrichs,
- Ralf Helbig,
- Julia Hilsenbeck,
- Prithvi Raj Pandey,
- Jens-Uwe Sommer,
- Lars David Renner,
- Tilo Pompe &
- …
- Carsten Werner

Show authors

Nature **volume 618**, pages 733–739 (2023)

- 2786 Accesses

- 59 Altmetric

- Metrics details

Abstract

Control of adhesion is a striking feature of living matter that is of particular interest regarding technological translation[1,2,3]. We discovered that entropic

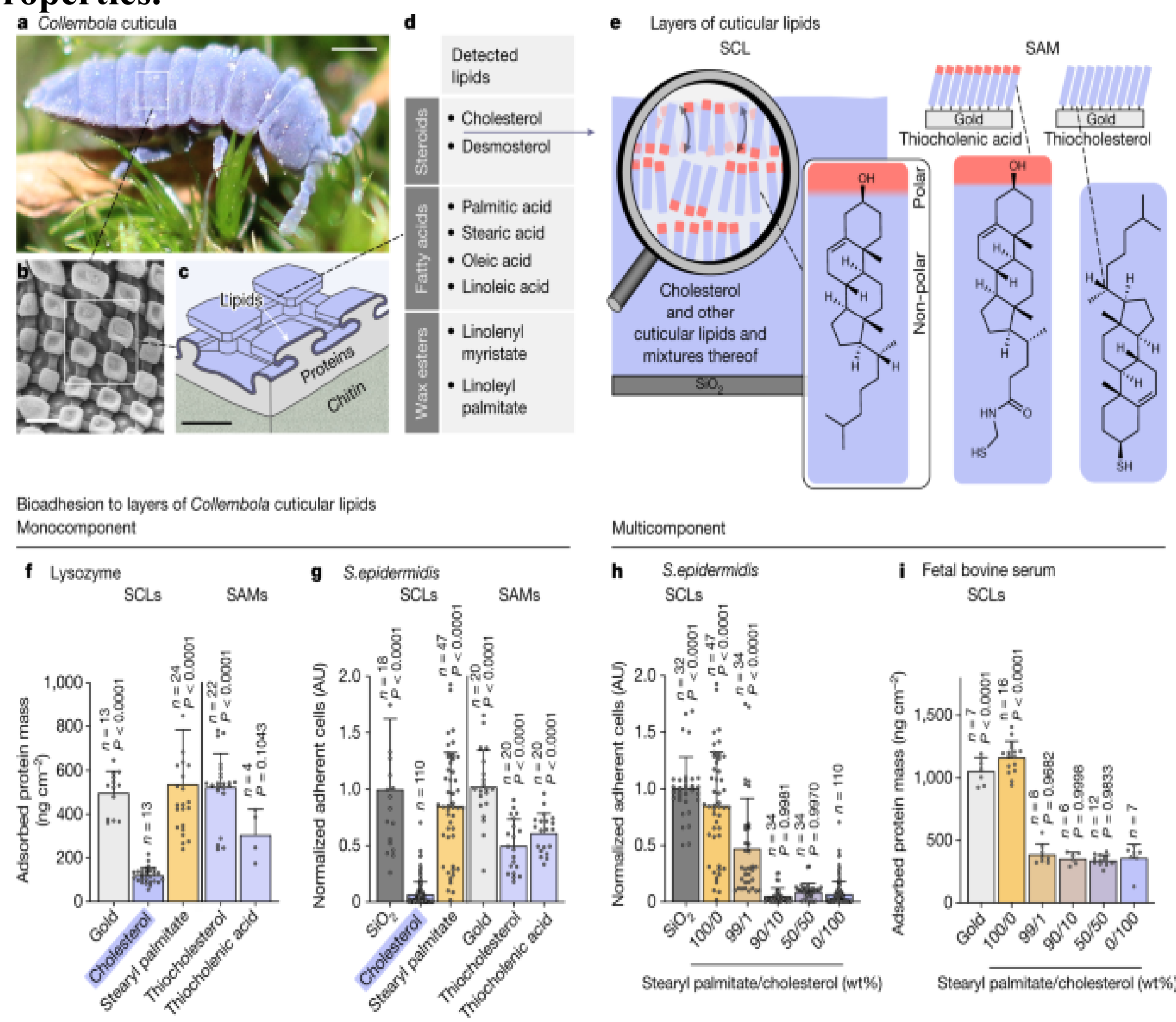
repulsion caused by interfacial orientational fluctuations of cholesterol layers restricts protein adsorption and bacterial adhesion. Moreover, we found that intrinsically adhesive wax ester layers become similarly antibioadhesive when containing small quantities (under 10 wt%) of cholesterol. Wetting, adsorption and adhesion experiments, as well as atomistic simulations, showed that repulsive characteristics depend on the specific molecular structure of cholesterol that encodes a finely balanced fluctuating reorientation at the interface of unconstrained supramolecular assemblies: layers of cholesterol analogues differing only in minute molecular variations showed markedly different interfacial mobility and no antiadhesive effects. Also, orientationally fixed cholesterol layers did not resist bioadhesion. Our insights provide a conceptually new physicochemical perspective on biointerfaces and may guide future material design in regulation of adhesion.

Main

Life has evolved a plethora of powerful principles to control adhesion, some of which have been recapitulated in engineered materials. Prominent examples include the superhydrophobic leaves of the sacred lotus[1] and the omniphobic surfaces of *Nepenthes* pitcher plants[2]. Whereas interfacial phenomena in nature are widely studied, the physical mechanisms underlying the control of bioadhesion—the interfacial accumulation of biopolymers and cells (including bacteria)—are not yet thoroughly understood. We previously explored the omniphobic, antibioadhesive cuticula of *Collembola* (Fig. 1a) and found that it consists of nanoscopic structures with overhanging cross-sectional profiles (Fig. 1b,c) preventing wetting and bacterial colonization[3,4,5,6]. Later the lipid-rich envelope of the *Collembola* cuticula (Fig. 1c)—considered to serve as another 'line of defence' against bioadhesion—was shown to contain aliphatic hydrocarbons, in particular steroids, fatty acids and wax esters (Fig. 1d and Extended Data Fig. 1)[7]. Whereas wax esters can be reasonably assumed to support the non-wetting properties of the cuticula[8], the role of steroids and fatty acids remains elusive. Free fatty acids have been reported to kill or inhibit the growth of bacteria and fungi[9,10], and steroids have been found to reduce bioadhesion on sponges and sea stars[11]; however, no mechanistic

explanation for this effect of steroids is available. Amphiphilic lipid components of the *Collembola* cuticula are also contained in the membranes of animal and bacterial cells that play key roles in compartmentalization and the functional alignment of molecular machineries[12]. Cholesterol, in particular, has been comprehensively studied and its presence is considered crucial for the regulation of functional lipid domains and the interaction between proteins and lipids[13]. However, the functional relevance of cholesterol at interfaces of living structures other than cellular membranes is underexplored.

Fig. 1: Layers of *Collembola* cuticular lipids and their bioadhesion properties.



a, Image of *Tetrodontophora bielanensis*, an exemplary *Collembola* sp. Scale bar, 1 mm. **b**, Scanning electron microscopy image of a *T. bielanensis*

cuticula. Scale bar, 500 nm. **c**, Cross-sectional schematic of the cuticula, showing a layered structure consisting of a chitin-rich inner skeleton covered by a protein-rich layer. A thin, lipid-rich envelope covers the protein-rich layer. Scale bar, 200 nm. **d**, Summary of lipids detected in the outer cuticula layer of *T. bielanensis*[7]. **e**, Layers of *Collembola* cuticular lipids; SCLs containing cholesterol facilitate orientational adaptation of the topmost lipids to the polarity of the environment. ATR–FTIR (Supplementary Fig. 2) and dynamic contact angle measurements (Fig. 2c and Extended Data Fig. 4a) indicate highly ordered cholesterol molecules, with the hydrocarbon tail of the outer cholesterol layer initially oriented towards the interface and the hydroxyl groups oriented inward. SAMs chemisorbed to gold via thiol groups, with either the polar or non-polar side of cholesterol oriented to the interface, served as references in selected experiments. **f**–**i**, Adsorbed amount of protein (**f**,**i**) and normalized adherent cells (**g**,**h**). Adsorbed amount of protein (lysozyme or fetal bovine serum) on monocomponent layers of *Collembola* cuticular lipids (**f**) and multicomponent SCLs of stearyl palmitate and cholesterol (**i**), as determined by quartz crystal microbalance measurements. Normalized adherent cells of *S. epidermidis* on monocomponent layers of *Collembola* cuticular lipids (**g**) and multicomponent SCLs of stearyl palmitate and cholesterol (**h**). Data normalized to average adherent cell density on a silica (SiO_2) substrate. **h**,**i**, Pure stearyl palmitate SCLs (100/0) and pure cholesterol SCLs (0/100) served as negative and positive controls, respectively. **f**–**i**, Mean + s.d. The number of observations (*n*) is indicated. *P* values (comparison with cholesterol SCL condition in **f**,**g** and with the 0/100 condition in **h**,**i**) were determined using one-way analysis of variance. AU, arbitrary units.

Source data

Using lipid layers without the cuticula's morphological features, we dissect the role of molecular assemblies in the control of bioadhesion from the antiadhesive topography effects of the *Collembola* cuticula[5,14]. Cholesterol-containing lipid layers were found to counteract bioadhesion by a previously unknown entropic repulsion mechanism that is—unlike earlier reported interfacial effects[15,16,17]—caused by orientational fluctuations.

Cholesterol layers exibit low bioadhesion

Lipids of the *Collembola* cuticula (Fig. 1d) were physiosorbed in multilayers on solid supports by spin-coating (spin-coated lipid multilayers, SCLs; Fig. 1e). For cholesterol SCLs, attenuated total reflection Fourier-transform infrared spectroscopy (ATR–FTIR) showed a multilayer structure with the majority of molecules oriented perpendicular to the interface (Supplementary Fig. 2)[18,19]. Self-assembled monolayers (SAMs; Fig. 1e) chemisorbed on gold served as a comparative analysis of the investigated lipidic cuticle components in geometrically constrained molecular orientations (Extended Data Figs. 1 and 2a–c).

A hallmark of bioadhesion is the formation of a molecular conditioning layer of proteins, nucleic acids, lipids and other biomolecules[20]. This process was mimicked by adsorption experiments with lysozyme and albumin, the results showing significantly lower adsorbed amounts of proteins on cholesterol SCLs than on all other studied SCLs and SAMs (Fig. 1f and Extended Data Fig. 2f,g). Likewise, adhesion of Gram-positive (*Staphylococcus epidermidis*, strain PCI 1200) and Gram-negative (*Escherichia coli*, strain W3110) bacteria was found to be very low on cholesterol SCLs (Fig. 1g and Extended Data Fig. 2d,e). By analysis of cholesterol SCLs of varying layer thickness (Supplementary Fig. 3a), any influence of multilayer thickness on antiadhesive properties could be excluded. Notably, SAMs of thiocholesterol and *N*-(2-mercaptoethyl)-3-hydroxy-5-cholenic acid amide (thiocholenic acid)—that is, surrogates of chemically fixed cholesterol monolayers of opposing orientational molecular alignment; Fig.1e)—accumulated higher amounts of adsorbed protein and adherent bacteria than cholesterol SCLs (Fig. 1f,g and Extended Data Fig. 2d–g). This observation suggests that restriction of molecular mobility in layered cholesterol impedes antiadhesive properties. Partial dissolution of cholesterol SCLs was excluded as a cause of low bioadhesion by total organic carbon content analysis (Supplementary Fig. 4).

Because cholesterol occurs in the *Collembola* cuticula together with several other lipid components (Fig. 1d), we further investigated its relevance regarding the adhesive characteristics of multicomponent SCLs. Multicomponent SCLs containing stearyl palmitate—a lipid that causes strong bioadhesion to its single-component SCLs (Fig. 1f,g and Extended Data Fig. 2d,g)—and cholesterol, with cholesterol content ranging from 0 to 100 wt%, were investigated using similar bioadhesion assays (Extended

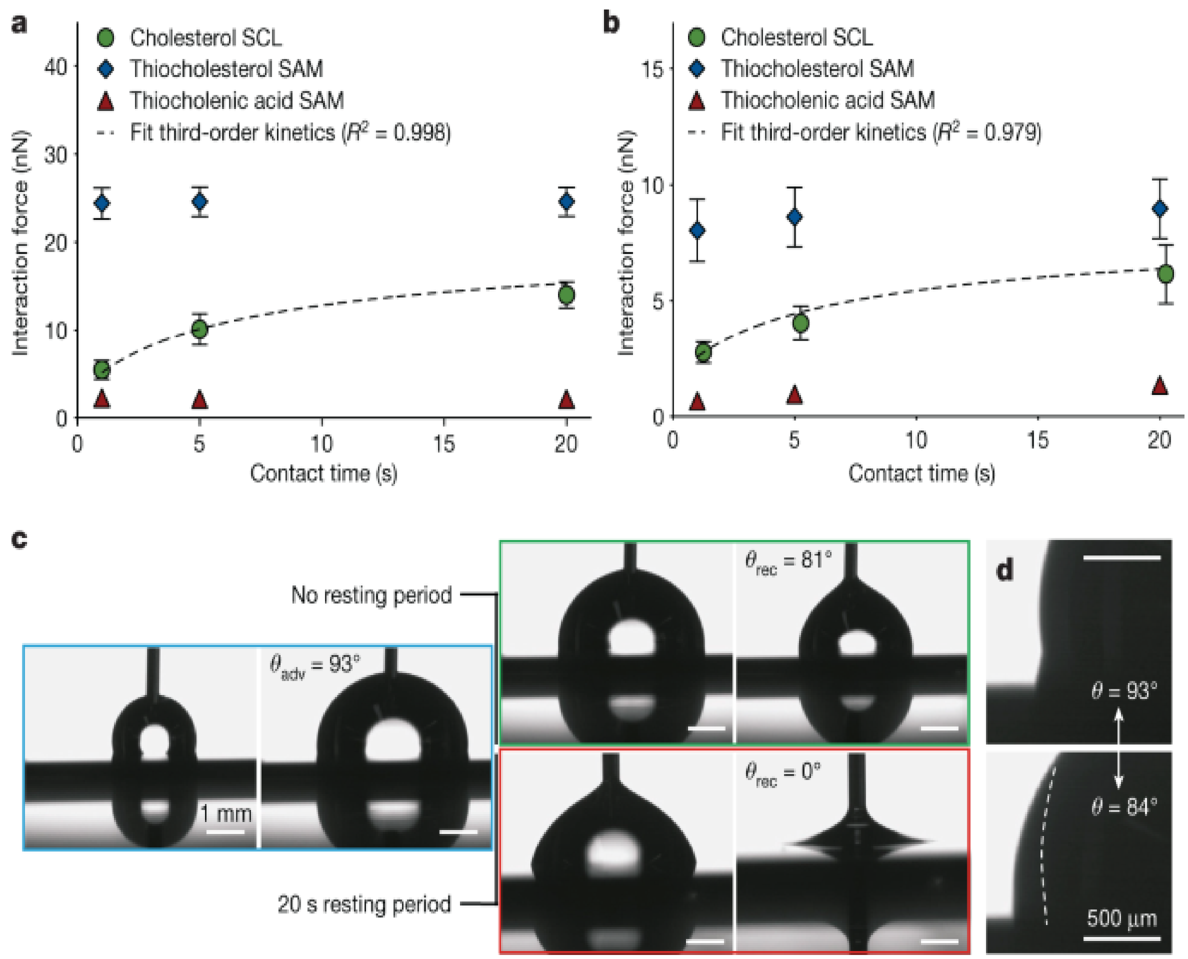
Data Fig. 3a–c). Multicomponent SCLs containing at least 10 wt% cholesterol showed low adherent cell densities similar to pure cholesterol SCLs (Fig. 1h and Extended Data Fig. 3d). Likewise, protein adsorption on multicomponent SCLs was significantly reduced when a minimum of only 1 wt% cholesterol was present (Extended Data Fig. 3e,f). The results of single-protein adsorption experiments were confirmed when applying physiologically relevant protein mixtures (that is, 10 vol% fetal bovine serum; Fig. 1i).

Cholesterol layers adapt to ambient polarity

The results of the bioadhesion assays suggest that mobility of interfacial cholesterol is key to the antiadhesive characteristics of single and multicomponent SCLs. Atomic force microscopy (AFM)-based colloid probe force spectroscopy, AFM-based single-cell force spectroscopy and dynamic contact angle measurements provided further insight into the underlying interfacial dynamics.

AFM-based force spectroscopy (Supplementary Fig. 5a,b and Methods) showed the dynamic interfacial adaptation of immersed cholesterol SCLs. Interaction forces between hydrophobic colloidal probes (Supplementary Fig. 5c) or single *E. coli* cells and cholesterol SCLs were quantified (Supplementary Fig. 5d) and were found to increase continuously with contact time (Fig. 2a,b). Further analysis showed that third-order kinetics controlled the underlying polarity adaptation process of the SCL interface to maximize hydrophobic interactions (Fig. 2a,b and Supplementary Note 1). Control experiments with thiocholesterol and thiocholenic acid SAMs confirmed that cholesterol mobility within SCLs is key to the observed adaptation (Fig. 2a,b).

Fig. 2: Cholesterol SCLs dynamically adapt to the polarity of their environment.



a,**b**, Contact time-dependent interaction forces (determined by AFM-based force spectroscopy) between a hydrophobic colloidal probe (**a**) (∅10 µm silica bead modified with a hydrophobic silane) or individual *E. coli* cells (**b**) (attached to a ∅10 µm silica bead) and cholesterol SCLs or thiocholesterol and thiocholenic acid SAMs. Experiments were performed at 37 °C. **a**,**b**, Mean ± s.e.m. Data were obtained from at least three independent experiments. Regressions are the best-fit solutions for a third-order process indicating formation of bonds at the probe cell–SCL interface following $\frac{{\rm{d}}\sigma }{{\rm{d}}t}=-\,k{\sigma }^{3}$ (for details see Supplementary Note 1). The interaction force was assumed to be directly correlated to the surface concentration of formed bonds, σ. Coefficients of determination (R^2) for fits are shown. **c**, Representative images showing the dynamic contact angles of water on the surface of a cholesterol SCL. The high advancing contact angle (θ_{adv}, blue box) reflects a hydrophobic surface. Following immediate withdrawal of the droplet (no resting period, green

box), a slightly reduced but still high receding contact angle (θ_{rec}) was observed. No receding of the three-phase contact line was observed when the water droplet was withdrawn after a 20 s resting period (red box). **d**, During the 20 s resting period, oscillations in the three-phase contact line between the initial advancing angle (93°) and an approximately 10° lower angle were observed.

Source data

Dynamic wetting experiments on cholesterol SCLs in air showed high advancing and receding water contact angles when the droplet was immediately withdrawn after deposition (Fig. 2c, Extended Data Fig. 4a and Supplementary Video 1), indicating that the hydrocarbon chains of the outer cholesterol layers conditioned in air are preferentially oriented towards the interface. If the water droplet was kept on the surface for 20 s before receding, oscillations of the three-phase contact line occurred for 15–20 s and subsequently the static contact angle was lower by approximately 10° (Fig. 2d and Supplementary Video 2). Pinning of the three-phase contact line was observed when the droplet was withdrawn (Fig. 2c, Extended Data Fig. 4a and Supplementary Video 2). The contact angle drop was found to be fully reversible following drying in inert gas (Extended Data Fig. 4d). The resting-period-dependent decrease in the contact angle, the oscillating three-phase contact line during the resting period and the increased contact angle hysteresis confirm a dynamic polarity adaptation of cholesterol at the interface of the SCLs in response to changes in the polarity of the environment.

The time dependence of the increase in interaction force between cholesterol SCLs and AFM probes (Fig. 2a) and single bacterial cells (Fig. 2b), as well as the decrease in receding water contact angle on cholesterol SCLs (Fig. 2c), indicate that the interfacial polarity adaptation process of cholesterol SCLs is relatively slow (that is, occurring over a period of a few seconds).

Also underpinning the results of the bioadhesion assays (Fig. 1h,i and Extended Data Fig. 3d–f), low quantities of cholesterol in multicomponent SCLs of cholesterol and stearyl palmitate sufficed to decrease the receding water contact angle (Extended Data Fig. 4c) and to increase the interaction force with hydrophobic AFM probes (Extended Data Fig. 3h).

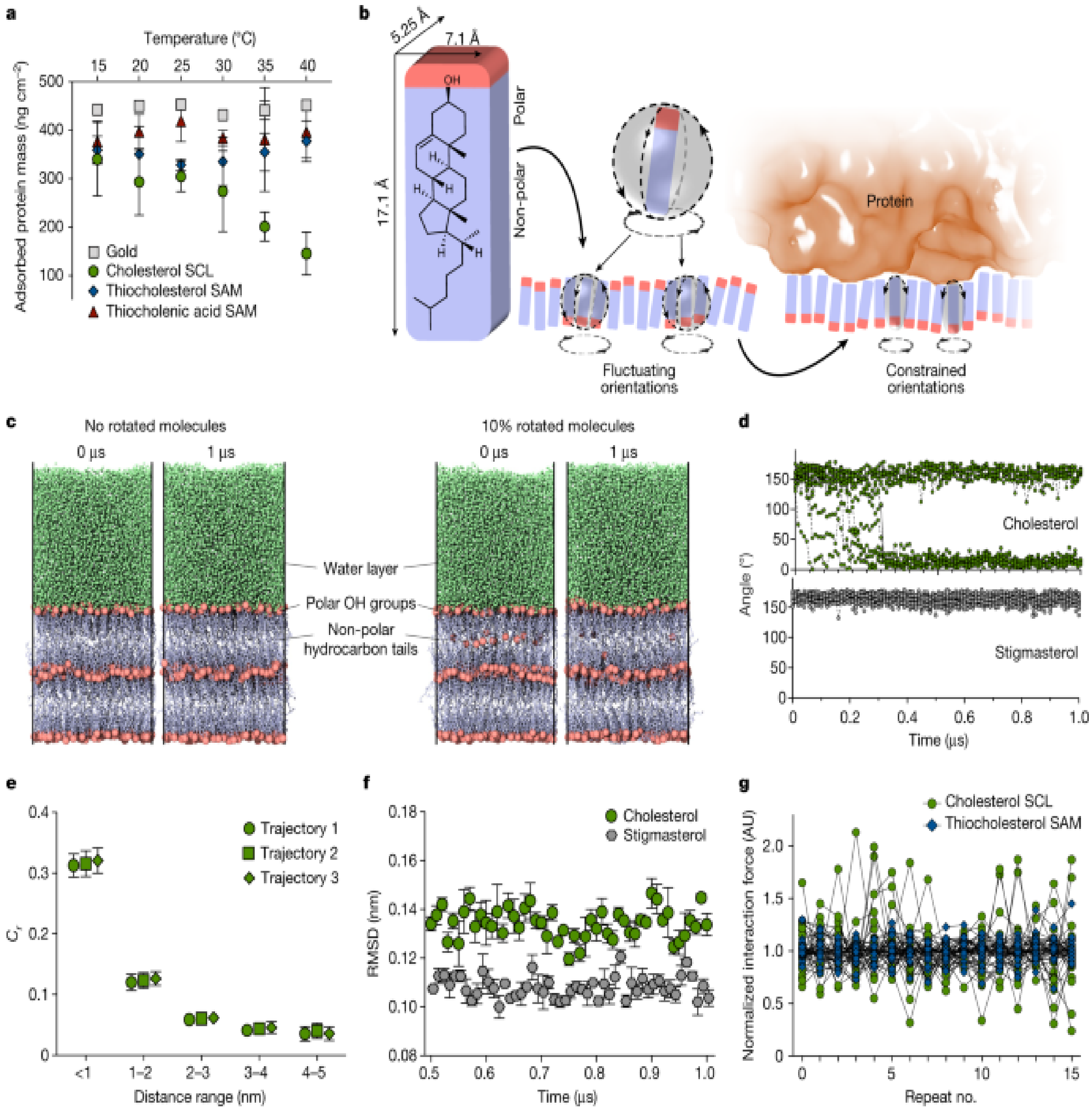
Entropic repulsion by surface fluctuations

The antibioadhesive properties of cholesterol-containing SCLs were shown above to correlate with a dynamic adaptation of the interfacial orientation of cholesterol in response to changes in environmental polarity. We hypothesized that entropically driven orientational fluctuations of interfacial cholesterol molecules mechanistically connect these features: any adsorption of biomolecules or attachment of (bacterial) cells requires an orientational (polarity) adaptation of the SCL interface that constrains the orientational states of interfacial cholesterol and thereby reduces the entropy of the system.

To validate this hypothesis we first examined the system for the characteristic temperature dependence of entropic effects on protein adsorption (according to the definition of the Gibbs free energy of adsorption, $\Delta G=\Delta H\,\mbox{--}\,T\Delta S$) using three selected proteins of increasing size and complexity: lysozyme, albumin and fibrinogen. Indeed, protein adsorption to cholesterol SCLs was observed to decrease when the temperature was increased from 15 to 40 °C (that is, in a temperature range in which significant temperature-induced changes in protein conformation can be excluded[21,22,23]), in contrast to minor changes in similar experiments with thiocholesterol and thiocholenic acid SAMs (orientationally fixed controls of cholesterol SCLs underneath the adsorbed protein) (Fig. 3a and Extended Data Fig. 5a,b). At thermal adsorption equilibrium, the temperature dependence of the Gibbs free energy of adsorption can be used to quantitatively estimate the repulsive entropic barrier (ΔS_{chol}) to protein adsorption on cholesterol SCLs. A quantitative analysis of adsorption (based on the slopes of $\Delta G(T)$ on cholesterol SCLs and the SAM controls) was performed for lysozyme, because adsorption-induced structural changes are negligible for the small, compact protein and provided $\Delta {S}_{{\rm{chol}}}=-\,200\,\pm \,60\,{\rm{J}}{{\rm{mol}}}^{-1}{{\rm{K}}}^{-1}$ (Extended Data Fig. 5c and Supplementary Note 2)[14,24]. This value of ΔS_{chol} shows a significant entropic repulsion counteracting bioadhesion to cholesterol SCLs. The plot of $\Delta G(T)$ for cholesterol SCLs exhibits a positive slope (that is, a negative ΔS of adsorption), indicating that entropic repulsion overcompensates any entropic gain of protein adsorption for that system. Whereas desorption-

determined processes might show a similar temperature dependence, protein adsorption is well known to be strongly determined by adsorption[24].

Fig. 3: Quantitative evidence of entropic repulsion by cholesterol SCLs and verification of the molecular mechanism.



a, Temperature-dependent adsorbed amounts of lysozyme as determined by quartz crystal microbalance measurements. For cholesterol SCLs, a pronounced temperature dependence on lysozyme adsorption was observed. Data were obtained from at least three independent experiments. Mean ± s.d. **b**, Schematic illustration of the transition of interfacial cholesterol from the

orientationally fluctuating (unbound) to orientationally constrained (protein-bound) state. Three aligned cholesterol molecules are shown to highlight the cooperative reorientation of neighbouring molecules within SCLs derived from geometric considerations and evidenced by the experimentally observed third-order kinetics of interfacial adaptation and orientational correlation length in MD simulations (Supplementary Notes 1 and 4). The entropic penalty of bioadhesion can be estimated from the ratio of constraint and fluctuating interfacial areas, $\Delta S = R \times \ln(A_{\mathrm{constr}}/A_{\mathrm{fluct}})$, of cholesterol molecules, with A_{fluct} roughly equal to the molecular dimension of cholesterol due to the cooperative orientational fluctuation and A_{constr} related to the cross-sectional area of the cholesterol facing the interface in the protein-bound state (Supplementary Note 3). **c**, MD simulations of cholesterol multilayers in contact with water, with equilibrium state (left) and intentionally reverted molecular orientation of 10% of molecules in interface layer (right) (for details see Supplementary Note 4). Oxygen atoms of hydroxyl groups are shown as orange spheres. **d**, Angles between the vector connecting atoms C3 and C17 on cholesterol and stigmasterol (Extended Data Fig. 6c and Supplementary Note 4) and the *z* axis for interfacial molecules as a function of simulation time for one representative simulation trajectory. In cholesterol multilayer systems with 10% of reverted molecules, spontaneous back-reorientation within 1 µs of simulation time was observed. No back-orientation was observed for stigmasterol controls. **e**, Correlation function (C_r) as a function of the distance to a reference molecule, plotted for three simulation trajectories of cholesterol multilayer systems (with no changed orientation of molecules at the interface). Error bars represent s.d. at that point over time intervals. **f**, Root mean squared deviation (RMSD) (for cholesterol multilayers and stigmasterol controls) over three simulation trajectories of the *z*-component of the position of all O atoms at the interface layer as a function of time for frames every 0.01 µs, with error bars denoting s.e.m. **g**, Normalized interaction forces between hydrophobic AFM tips and cholesterol SCL or thiocholesterol SAM surfaces. Each curve contains 16 data points recorded sequentially at the same location on the substrate. Individual datasets were acquired at different locations on the substrate. Curves are normalized to the average of each dataset. Data were obtained from at least three independent experiments.

Source data

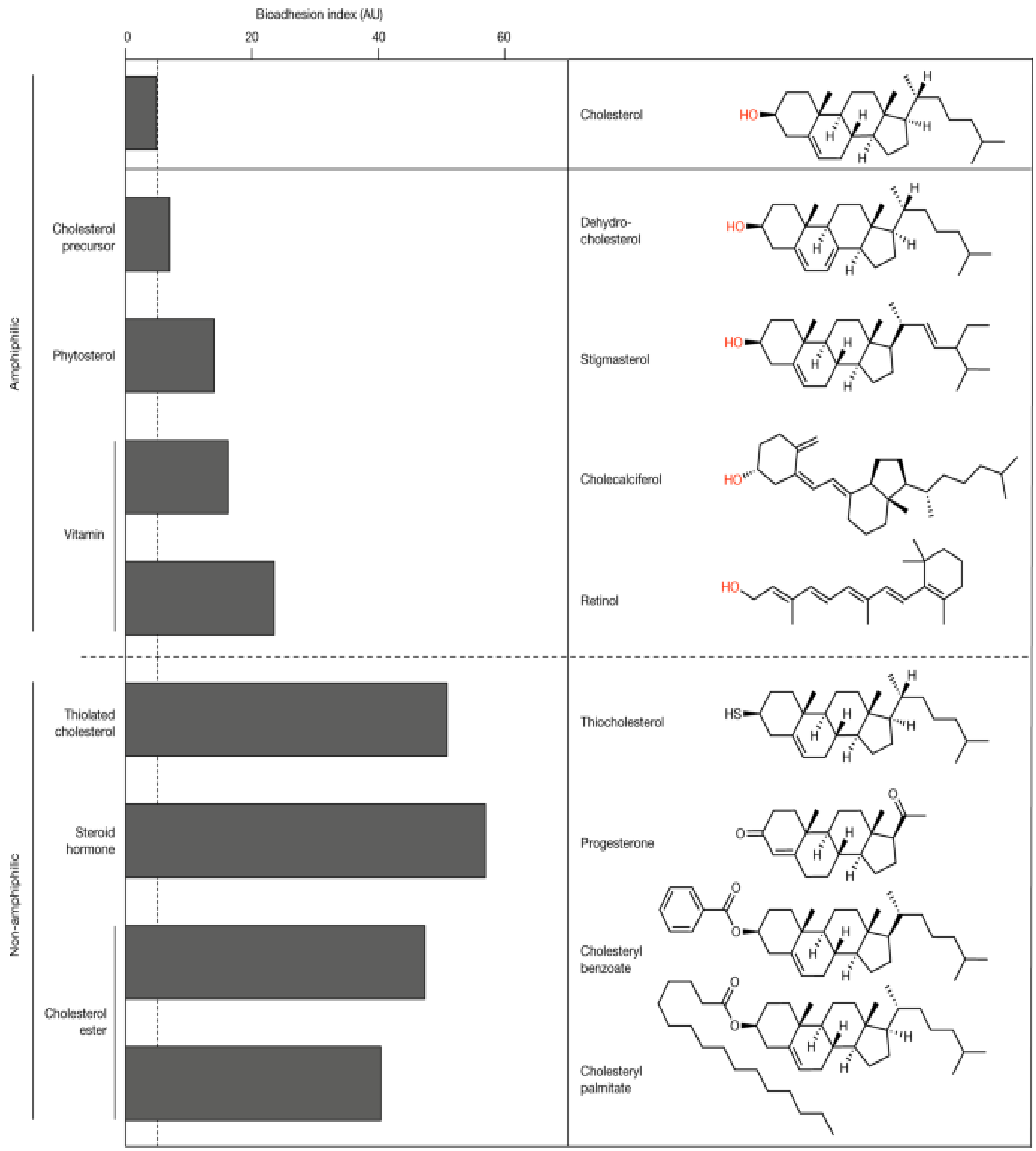
Entropic repulsion effects were previously reported to occur either due to the pinned conformational flexibility of grafted oligo(ethylene glycol)[15] and poly(ethylene glycol) brushes[16] or enforced smoothening of height fluctuations of lipid bilayers[17]. Herein, entropic repulsion of cholesterol SCLs is suggested to be mechanistically very different, primarily governed by polarity-controlled orientational fluctuations: The entropic penalty for the suppression of orientational fluctuations of cholesterol molecules at the SCL interface on protein adsorption can be estimated by relating the cross-sectional area of protein-facing cholesterol at the SCL interface to the area of orientationally free cholesterol at the SCL–solution interface (Fig. 3b and Supplementary Note 3)[18,19]. The obtained value for this entropic penalty (–160 J mol^{-1} K^{-1}) is close to the experimentally determined entropic repulsion barrier derived from temperature-dependent protein adsorption experiments (–200 J mol^{-1} K^{-1}). This rather large entropy penalty per adsorbing lysozyme molecule, corresponding to about three hydrogen bonds, emphasizes the relevance of the uncovered repulsive effect. In this estimation of entropic penalty, three associated interfacial cholesterol molecules were assumed to be cooperatively involved in orientational fluctuations, which was reasoned at first by geometrical considerations of cholesterol molecular packing in SCLs. Time-dependent interaction force measurements (Fig. 2a,b) showing a third-order kinetics support this assumption because the data indicate cooperative interactions involving three cholesterol molecules (Fig. 3b).

Analysis of adsorption experiments with the larger proteins albumin and fibrinogen yielded lower values for the estimated repulsive entropic barrier (probably due to entropy effects of conformational changes; Supplementary Note 2). However, it similarly showed a negative (repulsive) entropy contribution of cholesterol SCLs indicating the general relevance of the newly identified antiadhesion mechanism.

In an entirely independent approach, cholesterol multilayers in contact with water were explored by molecular dynamics (MD) simulations and similarly showed the relevance of molecular orientational fluctuations (Supplementary Note 4). Using equilibrium multilayer systems and steered MD simulations with reverted orientation of a variable percentage of interfacial cholesterol molecules, we observed spontaneous reorientation of these molecules within 1 μs simulation time such that their hydroxyl groups

aligned again toward the water layer (Fig. 3c,d and Extended Data Fig. 6b,d). Importantly, control simulations using the cholesterol analogue stigmasterol (Fig. 4) did not show this spontaneous molecular reorientation (Fig. 3c,d and Extended Data Fig. 6b,d). The orientation correlation function of cholesterol molecules in the interfacial layer indicated a rapid decay over a distance of 1 nm (Fig. 3e), confirming orientational cooperativity of a few (two or three) cholesterol molecules as discussed above. Together with the observation of high vertical fluctuation of the interfacial layer of cholesterol multilayers (Fig. 3f), the MD simulations support the view that cholesterol SCLs represent a highly dynamic and fluctuating molecular interface.

Fig. 4: SCLs of cholesterol analogue molecules show graded bioadhesion.



Bioadhesion to SCLs of cholesterol analogues (polar hydroxyl groups shown in red). To facilitate the comparison, a bioadhesion index was introduced by aggregating the results of bacterial adhesion experiments with *E. coli* and *S. epidermidis* cells and protein adsorption experiments with bovine serum albumin, lysozyme and fetal bovine serum (for the full set of experimental results see Extended Data Fig. 10), all normalized to the corresponding results for cholesterol SCLs. The reference value was calculated using the

respective controls from bacterial adhesion (SiO_2) and protein adsorption tests (gold). AU, arbitrary units.

Source data

To obtain direct experimental evidence of orientational fluctuations of cholesterol at SCL interfaces, we applied AFM-based force mapping to detect the simultaneous exposure of polar and non-polar residues of immersed cholesterol SCLs (Extended Data Fig. 7). The low and high levels of interaction forces between a hydrophobic AFM tip and thiocholenic acid or thiocholesterol SAMs span the range of interaction forces detected for cholesterol SCLs, confirming that both polar and non-polar residues are exposed at the interface of immersed cholesterol SCLs (Extended Data Fig. 7c). Repeated force–distance measurements at a constant position showed substantially higher variations for cholesterol SCLs than for thiocholesterol SAMs, providing proof of orientational fluctuations of cholesterol at the SCL interface (Fig. 3g and Extended Data Fig. 8).

Bioadhesion to layers of cholesterol analogues

The discovered entropic barrier to bioadhesion resulting from orientational fluctuations at the interface of cholesterol-containing SCLs raises the question of what key structural features of cholesterol underlie this effect.

Therefore, we compared the bioadhesion characteristics of SCLs prepared from structurally related cholesterol analogues (Fig. 4 and Extended Data Fig. 9) in terms of bacterial adhesion and protein adsorption (Extended Data Fig. 10). The results confirmed that molecular amphiphilicity is essential—but not sufficient—for low bioadhesion (Fig. 4 shows the results of different assays aggregated into a bioadhesion index) and showed that even small deviations in the molecular structure of cholesterol can strongly diminish its antiadhesive performance. Only SCLs of dehydrocholesterol, the closest chemical relative of the tested cholesterol analogues, approached the bioadhesion index of cholesterol SCLs.

Our results show that cholesterol assembles in molecular arrangements that are capable of entropically restricting bioadhesion. The combination of

the spatially decoupled amphiphilicity of cholesterol and its effective alignment in multilayers, generating a slowly adaptive, cooperative interfacial orientational mobility of the assemblies, was identified to be the prerequisite of pronounced entropic repulsion.

The resting-time-dependent decrease in the receding water contact angle indicates that, among the set of cholesterol analogues compared, only cholesterol and dehydrocholesterol SCLs meet these criteria (Extended Data Fig. 9d). Fluorescence recovery after photobleaching (FRAP) measurements on cholesterol SCLs providing lateral diffusion coefficients of 0.3–0.4 $\mu m^2 s^{-1}$ further support the ordered, but slowly adaptive, characteristics of the system, comparable to smectic liquid crystals and high-cholesterol-content lipid membranes (Supplementary Fig. 6)[25,26].

In agreement with the experimental results, MD simulations show that SCLs of analogues with slight structural modifications with respect to cholesterol (Extended Data Fig. 9d) behave significantly differently, with no spontaneous back-rotations of reverted molecules within the simulation time (Fig. 3d and Extended Data Fig. 6b,d). The simulations also indicate that stigmasterol SCLs are subject to significantly lower interfacial fluctuations than cholesterol SCLs, suggesting that the specific intermolecular interactions of cholesterol provide high interfacial mobility and allow for spontaneous molecular fluctuations (Fig. 3f and Extended Data Fig. 6f). These findings strongly support the experimentally deduced criteria for amphiphilic molecules capable of generating entropically repulsive assemblies.

Exploring the compositional line of defence of the antiadhesive *Collembola* cuticula, we identified physicochemical features of cholesterol-containing layers that result in effective entropic repulsion. Given the widespread occurrence of cholesterol at tissue boundaries, our findings may shed new light on important and ubiquitous interfacial processes of living matter. The development of cholesterol-inspired synthetic and scalable materials to control bioadhesion is highly attractive but demands further analyses of the molecular requirements that govern this newly identified mechanism.

Methods

Materials

All chemicals (solvents, lipids, media, proteins and so on) used in this article were purchased from Sigma-Aldrich without further purification, except where stated otherwise.

Lipid layer preparation

SCLs

SCLs were prepared on silicon wafers (10×15 mm^2). The substrates were cleaned by immersion in a solution of deionized water, ammonia and hydrogen peroxide (volume fraction 5/1/1) for 15 min at 70 °C, rinsed repeatedly in Milli-Q water and then dried in a nitrogen stream. The cleaned substrates were immediately used for the preparation of SCLs by spin-coating. Compounds were dissolved in chloroform at a concentration of 2 wt% (unless stated otherwise), and spin-coating (LabSpin6, SÜSS MicroTec) was performed at 3,000 rpm s^{-1} for 30 s.

SAMs

Silicon wafers were cleaned as described for SCL preparation. Clean substrates were first coated with a 3 nm chromium adhesion layer and then a 50 nm gold layer by chemical vapour deposition performed at 5×10^{-5} mbar (Univex 300, Leybold). Gold substrates were cleaned by immersion in a solution of deionized water, ammonia and hydrogen peroxide (volume fraction 5/1/1) for 15 min at 70 °C, rinsed repeatedly in Milli-Q water and then dried in a nitrogen stream. To minimize defects, the gold substrates were used for SAM formation immediately after cleaning. Thiol compounds were dissolved in ethanol (analytically pure, p.a.) to a concentration of 1 mM. Before use, thiol solutions were sonicated for 5–10 min. Gold-coated substrates were additionally cleaned in ethanol (p.a.) for 30 min using an ultrasonic bath. Subsequently, samples were immersed in thiol solutions and incubated for 24 h to ensure complete assembly. Containers housing the solution and samples were filled with dry nitrogen and sealed to minimize oxygen exposure. After incubation, sample surfaces were rinsed for 15–20 s

with ethanol (p.a.), dried under nitrogen and used directly for further experiments.

Preparation of thiocholenic acid

For the reaction scheme see Supplementary Fig. 1. One equivalent of 3-acetoxy-5-cholenic acid (**A**) was dissolved in anhydrous dimethyl sulfoxide (DMSO) under an argon atmosphere. The solution was stirred on ice, and six equivalents of carbonyldiimidazole (**B**), dissolved in anhydrous DMSO, were added slowly to the solution. The reaction mixture was stored at 4 °C overnight for activation. A solution of 2.5 eq. cystamine (**D**) was added dropwise to the stirred, ice-cooled imidazole intermediate (**C**) and left to react overnight at room temperature. The product (**E**) was purified, freeze-dried, dissolved in ethanol and diluted in water before the addition of 1 M NaOH to obtain ester cleavage. After purification and lyophilization, the product (**F**) was again dissolved in ethanol and diluted in water. The disulfide bond was cleaved by the addition of a sixfold excess of tris(2-carboxyethyl)phosphine hydrochloride to cholenic acid in aqueous solution at neutral pH. The cleavage product (**G**) was purified and lyophilized (purity around 80%). For reaction/purity control and purification, reverse-phase high-performance liquid chromatography (RP–HPLC) with a linear gradient of acetonitrile in water with additive formic or trifluoroacetic acid was used. The analytical HPLC instrument (1260 Infinity II, Agilent) was equipped with a diode array detector (210 and 278 nm) and an electron spray ionization–time-of-flight (ESI–TOF) detector. The preparative HPLC instrument (1200 series, Agilent) used a diode array detector and a fraction collector in manual collection mode.

TOC measurements

The total organic carbon (TOC) content of solutions was analysed using a Sievers 5310C laboratory TOC analyser (GE Analytical Instruments) in accordance with the manufacturer’s specifications. Information on the preparation of samples is provided in Supplementary Fig. 4.

Bacterial adhesion assays

S. epidermidis (strain PCI 1200, ATCC) and *E. coli* (strain W3110) were grown overnight from single colonies in lysogeny broth (LB) at 37 °C and 200 rpm. Overnight cultures were centrifuged at 4,000*g* for 5 min. The supernatant was removed and the remaining pellet resuspended in LB. This washing step was repeated three times. Cell densities were adjusted to an optical density (OD_{600}) of 0.2 in fresh LB, and sample substrates were incubated in the bacterial solution for 1 h at 37 °C (no shaking). After incubation, adherent bacteria were fixed with 4% paraformaldehyde in PBS for 10 min, washed in fresh PBS and Milli-Q water and dried under nitrogen. Samples were sputter-coated with a 15 nm gold layer (SCD 050, Balzers) and imaged by scanning electron microscopy (SEM; XL30 ESEM-FEG, Philips/FEI) in high-vacuum mode at an acceleration voltage of 5 kV. For each sample, at least six images were acquired at random positions and cells were counted using the counting tool in Fiji[27]. The scale bars of SEM images were used to calibrate pixel width. The Fiji 'cell counter' plugin was used to count the number of cells in the calculated area (approximately 10^3 μm^2). The numbers counted were either normalized against the median value of the silicon reference for relative comparisons or scaled up to cells per square millimetre for absolute values.

QCM measurements

Quartz crystal microbalance (QCM) measurements were performed using a QCM-D model E4 (Biolin Scientific) equipped with a peristaltic pump system (IPC, Ismatec). Gold-coated quartz crystals (QSX301, Quantum Design) with a resonance frequency of 5 MHz were used for QCM measurements. SCLs and SAMs were prepared on QCM crystals as described above. All measurements were performed at a flow rate of 100 µl min^{-1}. Protein solutions (lysozyme, bovine serum albumin and fibrinogen (100 µg protein ml^{-1} PBS)) or 10 vol% fetal bovine serum (Merck) in PBS were adsorbed on the layered samples for 1 h and subsequently subjected to a desorption regime for 30 min with PBS. Frequency and dissipation shifts induced by the adsorbed proteins were recorded in real time at the third, fifth, seventh, ninth, 11th and 13th overtones (15, 25, 35, 45, 55 and 65 MHz, respectively). The mass of adsorbed protein was calculated using the Sauerbrey equation[28] with Q-Sense DFind software (Biolin Scientific).

Dynamic contact angle measurements

Contact angle measurements were performed using an OCA 30 optical contact angle measuring and contour analysis system equipped with a TPC 160 temperature-controlled chamber (DataPhysics Instruments). Droplets of degassed deionized water were dispensed and redispensed at varying velocities of 0.3–2.0 µl s^{-1} to monitor advancing and receding contact angles and their time-dependent behaviour.

Ellipsometry

Ellipsometry measurements were performed with an M-2000 ellipsometer (J. A. Woollam) equipped with a 50 W QTH lamp operating at wavelengths from 371 to 1,000 nm and an angle of incidence of 75°. The oxide layer thickness of the silicon wafer was determined by ellipsometry before layer assembly. The thickness of the assembled layers was calculated by an optical model that included three layers: Si, SiO_2 and a Cauchy layer.

In situ ATR–FTIR

For in situ ATR–FTIR, 500 µl of a 2% cholesterol-chloroform solution was spin-coated on a germanium ATR crystal. Characterization of the deposited cholesterol SCLs by in situ ATR–FTIR was performed as described previously[29]. In situ ATR–FTIR spectroscopy was conducted using the single-beam sample reference technique to obtain compensated ATR–FTIR spectra in dry and aqueous environments[30,31]. Dichroic measurements were performed according to a previously described method[32]. Infrared light was polarized by a wire grid polarizer (SPECAC). The ATR–FTIR attachment was operated on an IFS 55 Equinox spectrometer (BRUKER Optics) equipped with a Globar source and mercury-cadmium-telluride detector. P- and s-polarized spectra were recorded from dry cholesterol SCLs. The high dichroic ratios of the n(C-O) band of the C-O-H headgroup of cholesterol can be verified in the line of the ATR–FTIR dichroism measurements of lipid bilayers. The dichroic ratio $R = A_P/A_S$, with absorbance A_p measured in p-polarization and A_S measured in s-polarization, of infrared bands with a transition dipole moment (M) located perpendicular to the surface plane also

showed high values, with $R > 4$. Briefly, under conditions of ATR at the interface of the dense medium (Si) and rare medium (air), an evanescent wave was established with an electrical field split into the three electrical field components, E_x, E_y and E_z, which interact with, for example, adjacent organic layers. Parallel polarized infrared light (E_P) forms E_x and E_z whereas vertically polarized light (E_S) forms E_y. High values of either R or A_p are obtained when the M of a functional group within the organic layer lies parallel to E_z (out of plane), whereas low R or high A_S values are obtained when M lies parallel to E_y (in plane), which is due to the scalar product \(A=E\times M=E\times M\times \cos (E,M)\) of the vectors **E** and **M**.

Time-of-flight secondary ion mass spectrometry

Time-of-flight secondary ion mass spectrometry (ToF–SIMS) was conducted with a ToF–SIMS 5-100 instrument equipped with a 30 kV Bi liquid metal ion gun (IONTOF). Data were acquired in Bi_3^{++} mode and calibrated against a list of reference peaks (SurfaceLab7, IONTOF). The area of analysis was 300 × 300 μm^2, which was scanned over 128 × 128 pixels. The sampling depth of this technique is as low as a few nanometres—that is, only the uppermost molecular layers of the sample contribute to the analysis. Characteristic signals of cholesterol (mass to charge ratio, $m/z = 369.3$) and stearyl palmitate ($m/z = 257.2$) were selected for semiquantitative characterization of SCLs.

FRAP

We applied a previously described FRAP protocol to analyse diffusion coefficients and mobile fractions with a fluorescence confocal laser scanning microscope using the FRAP tool (SP5, Leica)[33,34]. For FRAP measurements, fluorescent cholesterol SCLs were prepared by the addition of 1/100 or 1/20 NBD cholesterol (ThermoFisher) to pure cholesterol solutions (2 wt%) and spin-coating to clean no. 1.5 glass coverslips (Corning). Cholesterol SCLs were subsequently submerged in deionized water or PBS, and FRAP was performed by photobleaching a defined spot with a diameter of 10 ± 1 µm using the following protocol: ten images before bleaching followed by a high-power laser beam with subsequent bleaching

for 4 s to achieve a completely bleached area. Recovery was recorded with a 40×/ 1.4 numerical aperture oil immersion objective at an image acquisition speed of 1 s per frame at 256 × 256 pixels, for a total of 300 s (SP5, Leica). The resulting time-lapse was analysed using the MATLAB (MathWorks) programme *frap_analysis*[35].

AFM

All AFM measurements were performed with a NanoWizard IV AFM (JPK Instruments). The cantilevers used were calibrated before measurements. Measurements were conducted in PBS at room temperature (25 °C), except where stated otherwise.

Topographic imaging

Surface topography of the layered surfaces was recorded using the Quantitative Imaging mode of the AFM instrument using qp-BioAC cantilevers (Nanosensors). The acquisition parameters employed were as follows: 300 nm ramp, 10 ms pixel time and force trigger of 100 pN. Images of 30 × 30 μm^2 with a resolution of 256 × 256 pixels were recorded. The data-processing software provided by the AFM manufacturer (JPK Instruments) was used to extract the surface roughness (R_a) from topography images.

Colloidal probe force spectroscopy

For colloidal probe measurements, individual silica beads (Kisker Biotech, Ø10 µm) were attached to a tipless cantilever (PNP-TR-TL-Au, Nanoworld, nominal force constant 0.08 $N\ m^{-1}$) as described previously[36]. Colloidal probe-modified AFM cantilevers were cleaned in isopropanol, and adsorbed water was removed by heating at 120 °C for 10 min. The colloidal probe was hydrophobized by incubation in hexamethyldisilazane vapour for 12 h and subsequent heating at 120 °C for 1 h. The force spectroscopy parameters employed were as follows: 3 nN setpoint force, 5 $\mu m\ s^{-1}$ approach/retract velocity and 5 µm pulling distance. Interaction forces were extracted from

the retraction force–distance curves using data from the processing software provided by the AFM manufacturer.

Single-cell force spectroscopy

The colloidal probe-modified AFM cantilever (see section 'Colloidal probe force spectroscopy') was made cell adhesive by application of a polydopamine coating, and individual *E. coli* cells with cytoplasmic green fluorescent protein (strain MG1655 eGFP) were attached as described previously[37]. Measurements were conducted at 37 °C using a PetriDishHeater (JPK Instruments). The same force spectroscopy parameters and data-processing routines were used for colloidal probe spectroscopy measurements. Only datasets in which the position and orientation of the bacterial cell on the AFM cantilever was unchanged before and after measurements (that is, the contact conditions/geometry were constant during the measurement) were analysed.

Force spectroscopy measurements to quantify the spatial heterogeneity of cholesterol SCLs

Measurements were performed using the Quantitative Imaging mode of the AFM instrument using qp-BioAC cantilevers (CB-2, Nanosensors). The cantilevers were either hydrophobized by incubation in hexamethyldisilazane vapour as described above or hydrophilized by plasma cleaning (Harrick Plasma) for 10 min. The acquisition parameters used were as follows: 100 nm ramp, 20 ms pixel time and force trigger of 500 pN. Images of 5 × 5 μm^2 with a resolution of 50 × 50 pixels were recorded at different locations on the sample surface. Interaction forces were extracted from retraction force–distance curves using the data-processing software provided by the AFM manufacturer.

Force spectroscopy measurements to quantify the temporal heterogeneity of cholesterol SCLs

Time-dependent force spectroscopy measurements were performed with hydrophobized (see previous section) qp-BioAC cantilevers (CB-2,

Nanosensors). The force spectroscopy parameters employed were as follows: 1 nN setpoint force, 1 µm s^{-1} approach/retract velocity and 200 nm pulling distance. A waiting period of 20 s was maintained between the 16 consecutive measurements made at a single location on the sample surface, to minimize the influence of measurements on the dynamics of cholesterol molecules. Measurements were repeated at different locations of the sample. Interaction forces were extracted from retraction force–distance curves using the data-processing software provided by the AFM manufacturer.

MD simulations

Cholesterol or stigmasterol SCLs with four molecular layers—that is, two double layers (Extended Data Fig. 6a)—were modelled. At the interface of two double layers the hydrophilic hydroxyl groups of cholesterol or stigmasterol face each other; at the interface, hydroxyl groups face the water layer. The double layer–water interface normal was along the *z* axis in all simulations. To model a solid substrate at the bottom of the SCL, the motion of molecules was restrained in the *xy* plane for the lowermost lipid molecules. The simulation box dimensions were 7.1554 × 7.1554 × 50 nm^3. Hydrophobic walls, modelled as direct 12-6 LJ potential at $z = 0$ and $z = 50$ nm, were included at the bottom and top of the simulation box along the *z* axis. Vacuum layers above and below the multilayer prevent water–wall and lipid–wall short-range non-bonded interactions.

Cholesterol and stigmasterol molecules were modelled with the CHARMM36 force field[38,39]. The TIP3P water model in CHARM[40,41,42] was used and included KCl salt in the simulations. First, double layers of cholesterol or stigmasterol were constructed on CHARMM-GUI, where the hydrophobic tails face each other[43,44]. The CHARMM36 force field parameters for cholesterol and stigmasterol, TIP3P water parameters and parameters for ions were obtained from CHARMM-GUI. Each layer in the double layer of cholesterol or stigmasterol contained 128 molecules. The double layer obtained from CHARMM-GUI was translated along the *z* axis by 4 nm using the Visual Molecular Dynamics visualization programme[45], constructing multilayers of four layers containing 512 lipid molecules in total (Extended Data Fig. 6a). Water and ions were added on top of the multilayers. The cholesterol multilayer system contained 12,284 water

molecules with 72 K^+ and Cl^- ions and the stigmasterol multilayer system contained 12,363 water molecules with 74 K^+ and Cl^- ions. The system was energy minimized and equilibration simulations for both cholesterol- and stigmasterol-containing systems were conducted using Gromacs 2019.4 (for details see Supplementary Note 4)[46,47].

To construct the systems with a reversed molecular orientation at the interface (top layer), 10% (13 molecules), 30% (39 molecules) and 50% (64 molecules) of the molecules were reversed in orientation compared with the equilibrium system. The Alchembed tool[48] was used to remove any overlaps between coordinates that might have appeared on reverting the orientation of molecules. Minimization and short equilibration runs were then conducted as described above (for details see Supplementary Note 4).

Data availability

All data generated during this study are provided at https://is.gd/3cdZ6f. Source data are provided with this paper.

References

1. Barthlott, W. & Neinhuis, C. Purity of the sacred lotus, or escape from contamination in biological surfaces. *Planta* **202**, 1–8 (1997).

2. Wong, T. S. et al. Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature* **477**, 443–447 (2011).

3. Hensel, R., Neinhuis, C. & Werner, C. The springtail cuticle as a blueprint for omniphobic surfaces. *Chem. Soc. Rev.* **45**, 323–341 (2016).

4. Hensel, R. et al. Biologically inspired omniphobic surfaces by reverse imprint lithography. *Adv. Mater.* **26**, 2029–2033 (2014).

5. Hensel, R. et al. Tunable nano-replication to explore the omniphobic characteristics of springtail skin. *NPG Asia Mater*. **5**, e37 (2013).

6. Helbig, R., Nickerl, J., Neinhuis, C. & Werner, C. Smart skin patterns protect springtails. *PLoS ONE* **6**, e25105 (2011).

7. Nickerl, J., Tsurkan, M., Hensel, R., Neinhuis, C. & Werner, C. The multi-layered protective cuticle of *Collembola*: a chemical analysis. *J. R. Soc. Interface* **11**, 20140619 (2014).

8. Koch, K. & Ensikat, H. J. The hydrophobic coatings of plant surfaces: epicuticular wax crystals and their morphologies, crystallinity and molecular self-assembly. *Micron* **39**, 759–772 (2008).

9. Pohl, C. H., Kock, J. L. F. & Thibane, V. S. in *Science Against Microbial Pathogens: Communicating Current Researchand Technological Advances* Vol. 1 (ed. Méndez-Vilas, A.) 61–71 (2011).

10. Desbois, A. P. & Smith, V. J. Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Appl. Microbiol. Biotechnol.* **85**, 1629–1642 (2010).

11. Guenther, J., Wright, A. D., Burns, K. & De Nys, R. Chemical antifouling defences of sea stars: effects of the natural products hexadecanoic acid, cholesterol, lathosterol and sitosterol. *Mar. Ecol. Prog. Ser.* **385**, 137–149 (2009).

12. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112–124 (2008).

13. Lingwood, D. & Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **327**, 46–50 (2010).

14. Hensel, R. et al. Wetting resistance at its topographical limit: the benefit of mushroom and serif T structures. *Langmuir* **29**, 1100–1112 (2013).

15. Li, L., Chen, S., Zheng, J., Ratner, B. D. & Jiang, S. Protein adsorption on oligo(ethylene glycol)-terminated alkanethiolate self-assembled monolayers: the molecular basis for nonfouling behavior. *J. Phys. Chem. B* **109**, 2934–2941 (2005).

16. Jeon, S. I., Lee, J. H., Andrade, J. D. & De Gennes, P. G. Protein–surface interactions in the presence of polyethylene oxide: I. Simplified theory. *J. Colloid Interface Sci.* **142**, 149–158 (1991).

17. Lipowsky, R. in *Handbook of Biological Physics* Vol. 1 (eds. Lipowsky, R. & Sackmann, E. B. T.-H.) 521–602 (North-Holland, 1995).

18. Loomis, C. R., Shipley, G. G. & Small, D. M. The phase behavior of hydrated cholesterol. *J. Lipid Res.* **20**, 525–535 (1979).

19. Shieh, H. S., Hoard, L. G. & Nordman, C. E. Crystal structure of anhydrous cholesterol. *Nature* **267**, 287–289 (1977).

20. Donlan, R. M. Biofilms: microbial life on surfaces. *Emerg. Infect. Dis.* **8**, 881 (2002).

21. Venkataramani, S., Truntzer, J. & Coleman, D. R. Thermal stability of high concentration lysozyme across varying pH: a Fourier transform infrared study. *J. Pharm. Bioallied Sci.* **5**, 148–153 (2013).

22. Borzova, V. A. et al. Kinetics of thermal denaturation andaggregation of bovine serum albumin. *PLoS ONE* **11**, e0153495 (2016).

23. Marx, G. et al. Heat denaturation of fibrinogen to develop a biomedical matrix. *J. Biomed. Mater. Res. B Appl. Biomater.* **84B**, 49–57 (2008).

24. Norde, W. Biocolloids and biosurfaces energy and entropy of protein adsorption. *J. Dispers. Sci. Technol.* **13**, 363–377 (1992).

25. Lockwood, N. A., Gupta, J. K. & Abbott, N. L. Self-assembly of amphiphiles, polymers and proteins at interfaces between thermotropic liquid crystals and aqueous phases. *Surf. Sci. Rep.* **63**, 255–293 (2008).

26. Crane, J. M. & Tamm, L. K. Role of cholesterol in the formation and nature of lipid rafts in planar and spherical model membranes. *Biophys. J.* **86**, 2965–2979 (2004).

27. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

28. Sauerbrey, G. Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung. *Z. Phys.* **155**, 206–222 (1959).

29. Müller, M., Ouyang, W. & Keßler, B. Dichroic ATR-FTIR spectroscopy on oriented α-helical poly(l-lysine) multilayered with polyanions. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **77**, 709–716 (2010).

30. Fringeli, U. P. in *Encyclopedia of Spectroscopy and Spectrometry* (ed. Lindon, J. C.) 94–109 (Academic Press, 1999); https://doi.org/10.1016/B978-0-12-374413-5.00104-4.

31. Müller, M. et al. An in-situ ATR-FTIR study on polyelectrolyte multilayer assemblies on solid surfaces and their susceptibility to fouling. *Macromol. Rapid Commun.* **19**, 333–336 (1998).

32. Müller, M. Orientation of α-helical poly(L-lysine) in consecutively adsorbed polyelectrolyte multilayers on texturized silicon substrates. *Biomacromolecules* **2**, 262–269 (2001).

33. Zimmermann, R. et al. Charging and structure of zwitterionic supported bilayer lipid membranes studied by streaming current measurements, fluorescence microscopy, and attenuated total reflection Fourier transform infrared spectroscopy. *Biointerphases* **4**, 1–6 (2009).

34. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* **16**, 1055–1069 (1976).

35. Jönsson, P., Jonsson, M. P., Tegenfeldt, J. O. & Höök, F. A method improving the accuracy of fluorescence recovery after photobleaching analysis. *Biophys. J.* **95**, 5334–5348 (2008).

36. Krieg, M. et al. Tensile forces govern germ-layer organization in zebrafish. *Nat. Cell Biol.* **10**, 429–436 (2008).

37. Beaussart, A. et al. Quantifying the forces guiding microbial cell adhesion using single-cell force spectroscopy. *Nat. Protoc.* **9**, 1049–1055 (2014).

38. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).

39. Lim, J. B., Rogaski, B. & Klauda, J. B. Update of the cholesterol force field parameters in CHARMM. *J. Phys. Chem. B* **116**, 203–210 (2012).

40. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

41. Durell, S. R., Brooks, B. R. & Ben-Naim, A. Solvent-induced forces between two hydrophilic groups. *J. Phys. Chem.* **98**, 2198–2202 (1994).

42. Neria, E., Fischer, S. & Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **105**, 1902–1921 (1996).

43. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

44. Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

45. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

46. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

47. Abraham, M. J. et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

48. Jefferys, E., Sands, Z. A., Shi, J., Sansom, M. S. P. & Fowler, P. W. Alchembed: a computational method for incorporating multiple

proteins into complex lipid geometries. *J. Chem. Theory Comput.* **11**, 2743–2754 (2015).

Acknowledgements

We thank S. W. Grill, P. Fratzl, D. J. Müller and K. Simons for their most helpful and encouraging comments on our study. We thank D. Hahn, M. Nitschke and M. Müller for synthesis of thiocholenic acid, for performing ToF–SIMS measurements and for conducting ATR–FTIR measurements, respectively. L.D.R. acknowledges support from the Volkswagen Foundation. We thank C. Neinhuis for drawing our attention to the intriguing features of the *Collembola* cuticle.

Funding

Open access funding provided by Leibniz-Institut für Polymerforschung Dresden e.V.

Author information

Authors and Affiliations

1. Institute of Biofunctional Polymer Materials, Leibniz Institute of Polymer Research Dresden, Dresden, Germany

   Jens Friedrichs, Ralf Helbig, Julia Hilsenbeck, Lars David Renner, Tilo Pompe & Carsten Werner

2. Institute of Theory of Polymers, Leibniz Institute of Polymer Research Dresden, Dresden, Germany

   Prithvi Raj Pandey & Jens-Uwe Sommer

3. Cluster of Excellence Physics of Life and Center of Regenerative Therapies Dresden, Technische Universität Dresden, Dresden, Germany

Jens-Uwe Sommer & Carsten Werner

4. Institute for Biochemistry, Leipzig University, Leipzig, Germany

Tilo Pompe

Contributions

J.F., R.H., J.H., T.P. and L.D.R. designed and performed experiments and analysed data. P.R.P. and J.-U.S. performed molecular dynamics simulations. J.F., R.H., J.H., T.P., J.-U.S., L.D.R. and C.W. wrote the manuscript. All authors read and approved the final version of the manuscript.

Corresponding author

Correspondence to Carsten Werner.

Ethics declarations

Competing interests

J.F., R.H., L.D.R., T.P., J.-U.S. and C.W. are listed as co-inventors on patent application no. 10 2022 104 237.5. The remaining authors declare no competing interests.

Peer review

Peer review information

Nature thanks Mingjie Liu, Jose Luis Toca-Herrera and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Overview of lipids used for layer preparation.

Molecular structures of lipids detected in the cuticula of the Collembola *T. bielanensis*[7] and lipids used in the preparation of SCLs and SAMs. SCLs were prepared from cholesterol, palmitic and stearic acid, and stearyl palmitate. Cholesterol and palmitic and stearic acid have been detected in the cuticula, and stearyl palmitate is a substitute for the detected compound linoleyl palmitate, which was not available in sufficient quantities as a synthetic material. While both wax esters have the same hydrocarbon chain lengths, linoleyl palmitate has two additional double bonds. Other lipids detected in the cuticula were either not available as a synthetic material or unsuitable for the preparation of SCLs (i.e. oleic and linoleic acid are liquids at room temperature and immediately dewetted on the substrate after spin-coating). SAMs were prepared using thiol-terminated lipids structurally similar to some of the lipids detected in the Collembola cuticula. In thiocholesterol, the hydroxyl group at the 3rd carbon atom of hydrocarbon ring A of cholesterol is replaced by a thiol group and in cholenic acid, the cholesterol structure is chemically modified to carry a thiol group at the end of the hydrocarbon chain attached to the 17th carbon atom of hydrocarbon ring D (for details on the thiolation of cholenic acid, see Supplementary Fig. 1). Hexadecanethiol, octadecanethiol, and octadecenethiol are analogs of fatty acids detected in the cuticula, whereas thiohexyl hexanoate represents a short-chain substitute for stearyl palmitate.

Extended Data Fig. 2 Intrinsic characteristics and bioadhesion of *Collembola* cuticular lipid layers.

(**a**) Layer thickness, (**b**) surface roughness (R_a), and **(c)** morphology of *Collembola* cuticular lipid layers as determined by ellipsometry (layer thickness) and atomic force microscopy (surface roughness and

morphology). All images in panel (c) have the same scale. (**d,e**) Normalized adherent cell density of (**d**) *S. epidermidis* and (**e**) *E. coli*. Data are normalized to the average adherent cell density on the SiO_2 substrate. (**f,g**) Adsorbed amount of (**f**) lysozyme and (**g**) bovine serum albumin as determined by quartz crystal microbalance measurements. In all graphs, the mean + standard deviation are shown. The number of observations (n) *is* indicated. *P* values (comparisons to the cholesterol SCL condition) were determined using one-way ANOVA tests.

Source data

Extended Data Fig. 3 Bioadhesion characteristics of multicomponent SCLs.

(**a**) Layer thickness, (**b**) surface roughness (R_a), and (**c**) morphology of multicomponent SCLs containing stearyl palmitate and cholesterol in various weight ratios as determined by ellipsometry (layer thickness) and atomic force microscopy (surface roughness and morphology). All images in panel (c) have the same scale. (**d**) Normalized adherent cell density of *E. coli*. Data are normalized to the average adherent cell density on the SiO_2 substrate. (**e,f**) Adsorbed amount of (**e**) lysozyme and (**f**) bovine serum albumin as determined by quartz crystal microbalance measurements. The number of observations (n) is indicated. *P* values (comparisons to the condition "0/100") were determined using one-way ANOVA tests. (**g**) Characteristic ToF-SIMS signals of cholesterol (m/z 369.3) and stearyl palmitate (m/z 257.2), of multicomponent SCLs, as a function of layer composition. Rinsing the samples with Milli-Q water does not lead to significant changes in the detected signals, thus eliminating the possibility that the spatial arrangement of cholesterol and stearyl palmitate changes upon contact with polar fluids (i.e. interfacial enrichment of cholesterol on immersed multicomponent SCLs can be excluded as a reason for their anti-bioadhesive properties). In graphs (a-b, d-f), the mean + standard deviation are shown. **(h)** Contact time-dependent interaction forces (determined by AFM-based force spectroscopy) between a hydrophobic colloidal probe (∅10 µm silica bead modified with hydrophobic silane) and the surface of multicomponent SCLs. The mean and ± standard error of the mean are shown in graph (h).

Source data

Extended Data Fig. 4 Dynamic water contact angle measurements.

(**a–c**) Dynamic contact angle measurements on layers of Collembola cuticular lipids ((**a**) SCLs, (**b**) SAMs, and (**c**) multicomponent SCLs). For each sample type, the advancing (θ_{adv}) and receding (θ_{rec}) contact angle are displayed. For some samples the strong pinning of the three phase contact line after 20 s resting period (i.e. θ_{rec} is close to 0°) is additionally indicated by an arrow. Although palmitic and stearic acid are also amphiphilic, no resting-time-dependent decrease in the receding contact angle was observed for SCLs of either of these compounds. It can be assumed that the linear molecular structure of the fatty acids (in contrast to the bulky steroid structure of cholesterol) causes tighter packing of the molecules in the SCL, limiting the mobility of the molecules. In all graphs, the mean + standard deviation are shown. **(d)** Successive contact angle measurements at the same position on a cholesterol SCL, where the measuring site is dried with nitrogen between two measurements. Consistent results are obtained (i.e. a hydrophobic advancing contact angle and no receding angle after a 20 s resting period). Data were obtained from at least three independent experiments.

Source data

Extended Data Fig. 5 The temperature dependence of protein adsorption on cholesterol SCLs.

Temperature-dependent adsorbed amounts of (**a**) bovine serum albumin and (**b**) fibrinogen as determined by quartz crystal microbalance measurements. For cholesterol SCLs, a pronounced temperature dependence of protein adsorption is observed. Data were obtained from at least three independent experiments. The mean ± standard deviations are shown. (**c**) The temperature dependence of the Gibbs free energy of adsorption is derived from the amount of adsorbed lysozyme in quartz crystal microbalance measurements (Fig. 3a). Data points were calculated with $\Delta G = -RT \times \ln(q/(c^* - c^*q))$, where R is the universal gas constant, T the absolute

temperature, q the adsorbed surface fraction, and c^* the normalized protein concentration ($c^* = 6.8 \cdot 10^{-6}$). Surface fractions were calculated as the ratio of the experimentally determined amount to the maximum amount of saturated monolayers of adsorbed lysozyme (for detailed explanation see Supplementary Note 2). The mean ± standard errors derived from the protein adsorption measurements (at least three independent experiments) are shown. The coefficients of determination (R^2) for linear fits of the first-principle determination of the entropic barrier ΔS (by $\Delta G = \Delta H - T\,\Delta S$) are also given in the figure.

Extended Data Fig. 6 Molecular dynamics simulation of SCLs in contact with water.

(**a**) The left panel shows the simulation box for a system with four layers of cholesterol or stigmasterol in contact with a water layer and vacuum above and below. Hydrophobic walls were at z = 0 and z = 50 nm. The right panel shows the lipid multilayer–water system within the simulation box. For clarity, salt ions within the water layer are not shown. Oxygen atoms on the hydrophilic hydroxyl group of cholesterol or stigmasterol are shown as spheres and hydrophobic tails as blue sticks. (**b**) Simulation snapshots at 0 and 1 µs from a molecular dynamics (MD) trajectory for cholesterol and stigmasterol multilayer systems in contact with water with equilibrium states (0% rotated molecules) and intentionally reverted molecular orientation of 10, 30 and 50% of molecules in the interface layer. At 0 µs the production simulations were started after performing multiple equilibration steps (for details see Methods & Supplementary Note 4). (**c**) Visual representation of the vector connecting the C3 and C17 atoms on cholesterol or stigmasterol. This vector is used as the director for each molecule in the correlation analysis and the calculation of the angle of molecules at the interface layer (relative to the z-axis). (**d**) Angles between the vector connecting the C3 and C17 atoms and the z-axis from one representative simulation trajectory for lipid multilayer systems with 10, 30, and 50% reverted molecules. Cholesterol multiplayer systems show spontaneous reorientation within 1 µs of simulation time, whereas no reorientation is observed for stigmasterol controls. (**e**) Correlation function (C_r) as a function of the distance to a reference molecule, plotted for three simulation trajectories of stigmasterol multilayer systems (without changed orientation of molecules at the

interface), respectively. Error bars at each point represent the standard deviation at that point over the time intervals. (**f**) Histogram of mean RMSD over three simulation trajectories of z-component of the position of all oxygen atoms at the interface layer in the time range 0.5 to 1 µs shows higher fluctuation of the interface for cholesterol-containing multilayer system.

Source data

Extended Data Fig. 7 AFM-based force spectroscopy to probe the simultaneous exposure of polar and nonpolar molecular residues by cholesterol SCLs in a polar liquid.

(**a**) AFM-based, high-speed force mapping (*Quantitative Imaging*™, see Methods) was applied to detect the simultaneous exposure of polar (i.e., the cholesterol hydroxyl group) and nonpolar (i.e. the cholesterol hydrocarbon chain) molecular residues at the solid-liquid interface of cholesterol SCLs immersed in a polar liquid. Using polar (by applying an O_2 plasma in a plasma chamber) or nonpolar (hexamethyldisilazane coating) AFM tips (see (**b**)), interaction forces were quantified in a grid pattern on cholesterol SCLs, thiocholesterol SAMs (i.e., with the nonpolar cholesterol hydrocarbon chain permanently exposed), and thiocholenic acid SAMs (i.e. with the polar cholesterol hydroxyl group permanently exposed). The size ratios between an AFM tip and an individual cholesterol molecule are schematically shown. Although the AFM tip is very sharp (the typical radius of curvature of the tip is less than 10 nm), it is still significantly larger than a single cholesterol molecule. Thus, the individual force-distance curves record the integrated interactions of several cholesterol molecules. Force-distance curves were acquired at a high frequency (i.e., a few milliseconds per force-distance curve) that is sufficiently faster than the measured time for hydrophobic equilibration (i.e. a few seconds, Fig. 2c), thus capturing a 'snapshot' of the molecular surface orientation. (**c**) For nonpolar AFM tips, surface-dependent distributions of the interaction forces were found. For thiocholenic acid SAMs, predominantly low interaction forces were detected. The few higher forces may result from interactions of the AFM tip with the nonpolar part of the thiocholenic acid molecule, which can become accessible due to deformation of or defects in the SAM. Due to the

hydrophobic interactions between the nonpolar AFM tip and the exposed nonpolar hydrocarbon tail of the thiocholesterol molecule, significantly higher interaction forces were detected for thiocholesterol SAMs. For cholesterol SCLs, the interaction forces were observed to span the range of forces measured for the two SAMs studied, supporting the hypothesis that both polar and nonpolar residues are exposed on immersed cholesterol SCLs. (**d**) For polar AFM tips, no difference between the tested surfaces were detected. Data were obtained from at least three independent experiments.

Source data

Extended Data Fig. 8 AFM-based force spectroscopy to study the interfacial dynamics of cholesterol SCLs in a polar liquid.

(**a**) The interaction forces of a nonpolar AFM tip (see Extended Data Fig. 7b) with cholesterol SCLs and thiocholesterol SAMs were repeatedly quantified at a singular location on the sample surface. To minimize the influence of the AFM tip on the dynamics of the interfacial cholesterol molecules, the measurement parameters were chosen in such a way that a pause of 20 s was maintained between two consecutive measurements. Sixteen consecutive force-distance measurements were obtained per position with high spatial repeatability to ensure that an identical set of surface molecules contributed to the interaction with the AFM tip. At least six different locations per condition were analyzed on three independent sample surfaces. (**b**) Plot of the absolute interaction forces for all measurement series. The consecutive number of force-displacement curves measured at a single point and the elapsed time are plotted on the lower and upper x-axes, respectively. The individual measurement series were normalized to the mean value of each series. The resulting values are plotted in (**c**). (**d**) The distribution of the relative standard deviations determined from the individual curves in (c). The mean + standard deviation are shown. (**e**) Average slopes of linear regressions fitted to the individual measurement series presented in (c). There is only marginal evidence of a systematic change in interaction forces over the course of an individual measurement series (i.e., the recording of 16 force-distance curves at one position on the sample), whereby any influence of the measurement process due to wear or

contamination of the modified AFM tip or the sample surface can be excluded. The mean ± standard deviation are shown. The number of observations (n) is indicated. P values were determined using unpaired t-tests.

Source data

Extended Data Fig. 9 Characteristics of SCLs prepared from cholesterol analogs.

(**a**) Layer thickness, (**b**) surface roughness, and (**c**) morphology of cholesterol analog SCLs as determined by ellipsometry and atomic force microscopy. The contact mode atomic force microscopy images in (**c**) were used to determine the surface roughness. While all SCLs have a relatively uniform thickness of 150 – 200 nm, their morphology and roughness vary. However, a significant influence of the morphological surface properties on the bioadhesion properties is excluded, as e.g. cholesterol and dehydrocholesterol SCLs show a strongly diverging surface roughness but very similar bioadhesion properties (see Extended Data Fig. 10). (**d**) Dynamic contact angle measurements of the cholesterol analog SCLs. For SCLs of non-amphiphilic cholesterol analogs, only a surface roughness-dependent contact angle hysteresis (i.e. the rougher the surface, the higher the hysteresis due to pinning effects at elevations of the surface) but no resting time-dependent decrease of the receding contact angle (i.e. a dynamic interfacial polarity adaptation in response to changes in the polarity of the environment) is observed. Only for cholesterol and dehydrocholesterol SCLs a contact time-dependent decrease of the receding contact angle could be detected. In all graphs, the mean and + standard deviation are shown. The number of observations (n) is indicated.

Source data

Extended Data Fig. 10 Bioadhesion characteristics of cholesterol analog SCLs.

Normalized adherent cell density of (**a**) *S. epidermidis* and (**b**) *E. coli* on cholesterol analog SCLs. Data are normalized to the average adherent cell

density on the SiO_2 substrate. (**c,d,e**) Adsorbed amount of (**c**) lysozyme, (**d**) bovine serum albumin, and (**e**) 10% fetal bovine serum on cholesterol analog SCLs as determined by quartz crystal microbalance measurements. In all graphs, the mean + standard deviation are shown. The number of observations (n) is indicated. *P* values (comparison with the cholesterol condition) were determined using one-way ANOVA tests.

Source data

Supplementary information

Supplementary Information

This file contains Supplementary Figs. 1–8, Tables 1–3, Notes 1–4 and References.

Peer Review File

Supplementary Data

This zipped folder contains source data for Supplementary figures.

Supplementary Video 1

Video sequence of a dynamic water contact angle measurement on the surface of a cholesterol SCL. The high advancing contact angle reflects a hydrophobic surface. Following immediate withdrawal of the droplet (no resting period), a slightly reduced but still high receding contact angle is observed.

Supplementary Video 2

Video sequence of a dynamic water contact angle measurement on the surface of a cholesterol SCL. The high advancing contact angle reflects a hydrophobic surface. No receding of the three-phase contact line is observed when the water droplet is withdrawn after a rest period of 20 s. The water

droplet is then retracted. Oscillations in the three-phase contact line were observed during the 20 s rest period.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 9

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to

the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Friedrichs, J., Helbig, R., Hilsenbeck, J. *et al.* Entropic repulsion of cholesterol-containing layers counteracts bioadhesion. *Nature* **618**, 733–739 (2023). https://doi.org/10.1038/s41586-023-06033-4

- Received: 01 February 2021
- Accepted: 30 March 2023
- Published: 21 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06033-4

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Cholesterol can make surfaces non-stick

Nature Research Briefing 21 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06033-4

- Article
- Open Access
- Published: 31 May 2023

Increasingly negative tropical water–interannual CO_2 growth rate coupling

- Laibao Liu ,
- Philippe Ciais,
- Mengxi Wu,
- Ryan S. Padrón,
- Pierre Friedlingstein,
- Jonas Schwaab,
- Lukas Gudmundsson &
- …
- Sonia I. Seneviratne

Show authors

Nature **volume 618**, pages 755–760 (2023)

- 6200 Accesses

- 121 Altmetric

- Metrics details

Abstract

Terrestrial ecosystems have taken up about 32% of the total anthropogenic CO_2 emissions in the past six decades[1]. Large uncertainties in terrestrial carbon–climate feedbacks, however, make it difficult to predict how the land carbon sink will respond to future climate change[2]. Interannual variations in the atmospheric CO_2 growth rate (CGR) are dominated by land–atmosphere carbon fluxes in the tropics, providing an opportunity to explore land carbon–climate interactions[3,4,5,6]. It is thought that variations in CGR are largely controlled by temperature[7,8,9,10] but there is also evidence for a tight coupling between water availability and CGR[11]. Here, we use a record of global atmospheric CO_2, terrestrial water storage and precipitation data to investigate changes in the interannual relationship between tropical land climate conditions and CGR under a changing climate. We find that the interannual relationship between tropical water availability and CGR became increasingly negative during 1989–2018 compared to 1960–1989. This could be related to spatiotemporal changes in tropical water availability anomalies driven by shifts in El Niño/Southern Oscillation teleconnections, including declining spatial compensatory water effects[9]. We also demonstrate that most state-of-the-art coupled Earth System and Land Surface models do not reproduce the intensifying water–carbon coupling. Our results indicate that tropical water availability is increasingly controlling the interannual variability of the terrestrial carbon cycle and modulating tropical terrestrial carbon–climate feedbacks.

Main

The interannual variations (IAV) of the CO_2 growth rate (CGR) are found to be strongly correlated with El Niño/Southern Oscillation (ENSO)[12,13] (for example, $R = -0.55$, $P < 0.05$ in ref. 12, Pearson correlation coefficient), particularly with tropical temperature variations[7,8,9] (for example, $R = 0.7$, $P < 0.01$ in ref. 7), despite the lower IAV of tropical temperature than for other locations[14]. The historical IAV sensitivity of CGR to tropical temperature was further identified as an observational constraint that can significantly lower uncertainties in projected tropical carbon budgets[5]. Compared to tropical temperature, concurrent tropical precipitation is not well correlated with CGR[15,16] (for example, $R = -0.19$, $P > 0.1$ in ref. 16) but lagged tropical

precipitation was shown to strongly explain the IAV of CGR or tropical net land carbon flux[7,17](for example, $R = -0.5$, $P < 0.05$ in ref. [7]), resulting in an ambiguous role of water availability in controlling CGR from a process perspective. Recently, the launch of twin satellites of the Gravity Recovery and Climate Experiment (GRACE) enabled the direct measurement of terrestrial water storage (WS) variability, and a subsequent analysis showed that it is tightly coupled to CGR[11] ($R = -0.85$, $P < 0.01$). However, in the context of climate change, it remains unclear whether the identified terrestrial climate–carbon coupling is constant over time or may vary subject to changes in climate forcers and mean climate.

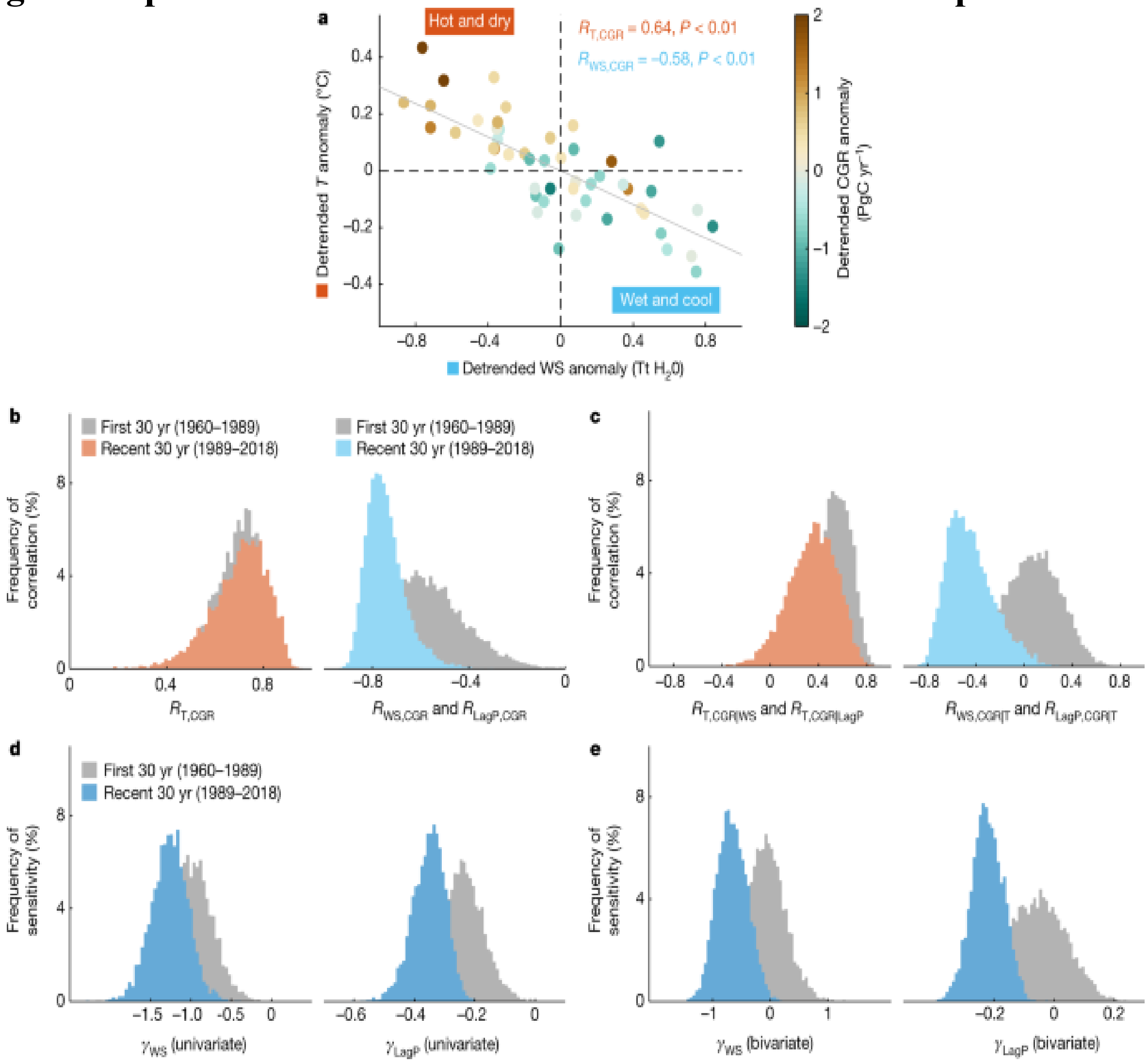
Here, we investigate the changes in the interannual relationship between tropical land climate conditions and CGR over the past decades. To complement the shorter observational record of the GRACE satellites, we also use recently reconstructed long-term WS variability[18]. Furthermore, the 6-month lagged yearly precipitation (LagP) can approximate aggregated tropical WS IAV well and correlates with CGR IAV, emerging as another efficient proxy for tropical terrestrial water availability IAV (Methods). This also helps to explain why lagged precipitation correlated well with CGR in previous findings[7,17].

Observed climate–carbon coupling

All variables are detrended at yearly time scale by removing the long-term linear trend, as we focus on the relationship in interannual variability. The years following the eruptions of Mount Agung (1962 and 1963), El Chichón (1982) and Mount Pinatubo (1991–1993) are also excluded from analyses to avoid perturbations of unusual carbon flux anomalies[19]. For the entire 1960–2018 period, CGR is significantly correlated with both tropical temperature ($R_{T,CGR} = 0.64$, $P < 0.01$, Pearson correlation coefficient) and tropical WS ($R_{WS,CGR} = -0.58$, $P < 0.01$) (Fig. 1a). The opposite sign in the two relationships suggests that hotter (positive temperature anomaly) and drier (negative WS anomaly) climate conditions generally dampened the land carbon sink and thus enhanced atmospheric CO_2 growth in the past decades. There is also a small proportion of CGR that does not fit the general pattern,

suggesting the role of other factors, such as exceptional (nonlinear) anthropogenic emissions or ocean carbon sink.

Fig. 1: Tropical land climate–carbon interannual relationships.



a, Yearly tropical temperature versus tropical WS versus CGR in detrended anomalies. The values of CGR are indicated by the colour bar. **b**, Histograms of climate–CGR interannual correlations in the first three decades (1960–1989) and in the more recent three decades (1989–2018), derived using 5,000 bootstrapping repeats. Both tropical WS and LagP are used to represent tropical water availability. **c**, As for **b** but showing histograms of the partial correlations of CGR to tropical temperature and tropical water after controlling tropical water and tropical temperature, respectively. **d**, Histograms of the interannual sensitivity of CGR to tropical

WS (γ_{WS}) and LagP (γ_{LagP}) in the univariate regression for the same two periods, derived using 5,000 bootstrapping repeats. Unlike correlations, γ_{WS} and γ_{LagP} differ in magnitude because of the differences in WS and LagP IAV magnitude and are therefore shown separately. The unit of this sensitivity is PgC yr^{-1} per Tt H_2O. **e**, As for **d** but showing γ_{WS} and γ_{LagP} estimated using the bivariate regression with both tropical water and tropical temperature as predictors. Ridge regression is used here to reduce biases from high collinearity between water and temperature (Methods).

Next, we investigate how the correlation between climate and CGR changes from the first 30 years (1960–1989) to the most recent 30 years (1989–2018). The uncertainty of the correlation is quantified through bootstrapping (5,000 replications). The results show that the bootstrap distributions of temperature–CGR correlations are similar between the two periods but water–CGR correlations are significantly different between the two periods (based on the Wilcoxon signed-rank test, $P < 0.05$) and become more negative over time (Fig. 1b and Extended Data Table 1). To check whether this increasingly negative water–carbon correlation is influenced by the possible confounding water–temperature coupling, we look at the temporal dynamic of water–temperature correlations and find that they are stable over time (Supplementary Fig. 1). Partial correlations remove water–temperature correlations and their relative changes directly help to confirm that the increasingly negative tropical water–carbon correlation remains robust (Fig. 1c). We note that because terrestrial water variability can also indirectly influence the land carbon cycle by triggering atmospheric temperature extremes through the well-documented soil moisture–atmosphere feedbacks[20,21], it could be inappropriate to interpret partial correlation $R_{W,CGR|T}$ as total water impacts on CGR in the two periods, but the temporal changes of $R_{W,CGR|T}$ are useful here (Methods). In addition, we compute climate–CGR correlations for a 25-year moving window to provide insights into the gradual changes, which are quite smooth over time (Supplementary Fig. 2). To further test the robustness of changes in the interannual correlations between tropical water and CGR, we also consider alternative observation datasets of tropical yearly precipitation and tropical temperature (Extended Data Fig. 1a). To verify that the IAV in CGR does not primarily originate from fossil fuel emissions, land-use change and ocean uptake, we

also use the residual land sink (RLS) instead of CGR (Extended Data Fig. 1b). These results all point to an increasingly negative correlation between tropical water and CGR on the interannual scale over the past decades.

The sensitivity of CGR IAV to tropical water is further estimated using linear regressions, defined as the slope of the regression between CGR and climate, with both variables detrended. In alignment with correlations, we first perform a univariate regression in which CGR is the function of tropical water alone to avoid possible underestimations of the sensitivity of CGR to tropical water variations. The univariate sensitivity of CGR to tropical WS and LagP increased (more negative) by about 35% on average from the previous 30 years (1960–1989) to the more recent 30 years (1989–2018) (Fig. 1d and Extended Data Table 1). For instance, the sensitivity to tropical WS increased from -0.95 ± 0.27 PgC yr^{-1} per Tt H_2O (teratonnes of water) to -1.26 ± 0.23 PgC yr^{-1} per Tt H_2O during the same two periods. We also perform bivariate linear regression with both tropical temperature and tropical water as predictors, and the sensitivity of CGR to tropical water variations is not significantly different from 0 in the first 30 years but becomes significantly negative in the recent 30 years ($P < 0.05$) (Fig. 1e). We use the ridge regression to reduce the effects of high collinearity between water and temperature on sensitivity estimates, but the ordinary least squares (OLS) regression also yields increasingly negative sensitivity over time (Extended Data Table 1). In addition, results based on a moving 25-yr window show that the sensitivity stalls within the recent 34-yr period, that is, after the time window centred on 1997 (1985–2009) (Supplementary Fig. 3). These results together suggest that the interannual relationship between tropical water and CGR has become increasingly negative in the recent past (1989–2018) compared to previous climate conditions (1960–1989). Also, we find that the amplitude of temperature sensitivity enhancement from the bivariate regression is smaller than that previously reported[16,22], if WS or LagP, rather than concurrent precipitation, is used as the proxy for water availability (Extended Data Fig. 2). This is because the accurate proxy for water availability was not identified and available (that is, using concurrent precipitation (reported $R_{Pre,CGR} = -0.19$, $P > 0.1$ in ref. [16]) rather than WS or LagP as the proxy starting from 1960 ($R_{LagP,CGR} = -0.68$, $P < 0.01$; Extended Data Fig. 3b)). Moreover, we extend the analysis time period by including 2011 to 2018 and observe a recent declining temperature sensitivity. All

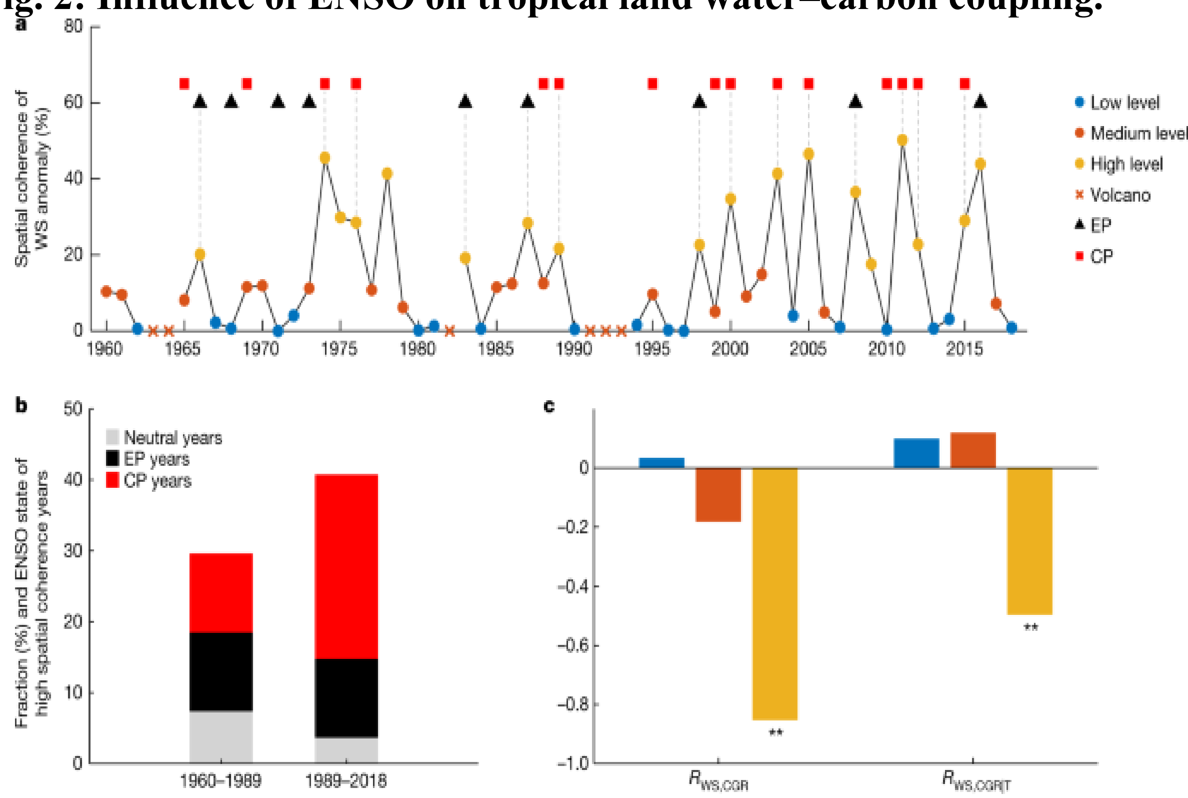
these results emphasize that it is crucial to integrate water availability into the carbon–climate feedback metric for better estimating climate-driven changes in tropical terrestrial carbon sink.

ENSO teleconnections

Interannual changes in tropical climate are largely driven by ENSO[23]. Under anthropogenic warming, growing observational evidence shows robust changes in ENSO characteristics, such as increases in frequency and variability with shifts in types[24], especially the largest sea surface temperature (SST) anomaly shifts from the Eastern Pacific (EP) to the Central Pacific (CP) since the 1990s[25,26,27,28]. As a result, it is likely to alter the patterns of moisture and heat fluxes over tropical continents, for instance, triggering more extreme droughts and fires[29] and thus modulating terrestrial carbon–climate feedbacks. Indeed, we find that most years with high spatial coherence of tropical WS anomalies are ENSO years (Fig. 2a,b), for which spatial coherence is quantified by applying a metric adjusted from ref. [9] (Methods). There is still a small proportion of high spatial coherence that cannot be explained by ENSO, suggesting a role for other factors, such as the tropical Atlantic variability[30,31] and the Indian Ocean Dipole[32,33]. Compared to ENSO-neutral years, spatial patterns of WS anomaly during ENSO years are quite uniform (Extended Data Fig. 4). Then, from the first 30 years (1960–1989) to the most recent 30 years (1989–2018), the fraction of years at high spatial coherence levels increases from 30% to 41% because of increased contributions of CP ENSO and fewer neutral years (Fig. 2b). These results could be important because the dominant role of water availability in controlling carbon fluxes at a larger scale is affected by the spatial compensatory degree of water anomalies: water availability clearly dominates carbon flux IAV locally but this control could partly be spatially counterbalanced when aggregated[9]. To further confirm the effects of spatial coherence on tropical water–CGR coupling, we perform a new subset analysis. We first detrend all years of data by removing the long-term trend and then bin them into three subsets according to spatial coherence levels. The $R_{\mathrm{WS,CGR}}$ is highly negative only when the spatial coherence is high (Fig. 2c). The $R_{\mathrm{WS,CGR|T}}$ confirms that this dependence of water–carbon coupling on spatial coherence is not influenced by confounding temperature effects.

Similar results are found when we replace WS with LagP (Supplementary Fig. 4). The sensitivity of CGR to WS is also expected to become apparently negative only when the spatial coherence is high (Supplementary Fig. 5). Another subset analysis using a 25-year moving window also supports the idea that the enhanced spatial coherence strengthens the increasingly negative tropical water–CGR coupling (Supplementary Fig. 6). These results indicate that identified ENSO-driven enhancement of spatial coherence over time is a likely explanation for the increasingly negative coupling between tropical water and CGR IAV. In the future, if CP ENSO events become more common with global warming, as predicted[34], tropical water might continue to increasingly control tropical terrestrial sink IAV. Further investigations on different impacts of EP ENSO and CP ENSO events on the terrestrial carbon cycle are useful to better understand future carbon–climate feedbacks[35,36].

Fig. 2: Influence of ENSO on tropical land water–carbon coupling.



a, Year-to-year variations of spatial coherence of tropical WS anomaly and ENSO. All years are classified into three subsets according to the level of spatial coherence: low level (0th to 33.3th percentile); medium level (33.3th to 66.6th percentile); and high level (66.6th to 100th percentile). Year is considered Eastern Pacific (EP) ENSO when the largest DJF SST anomaly

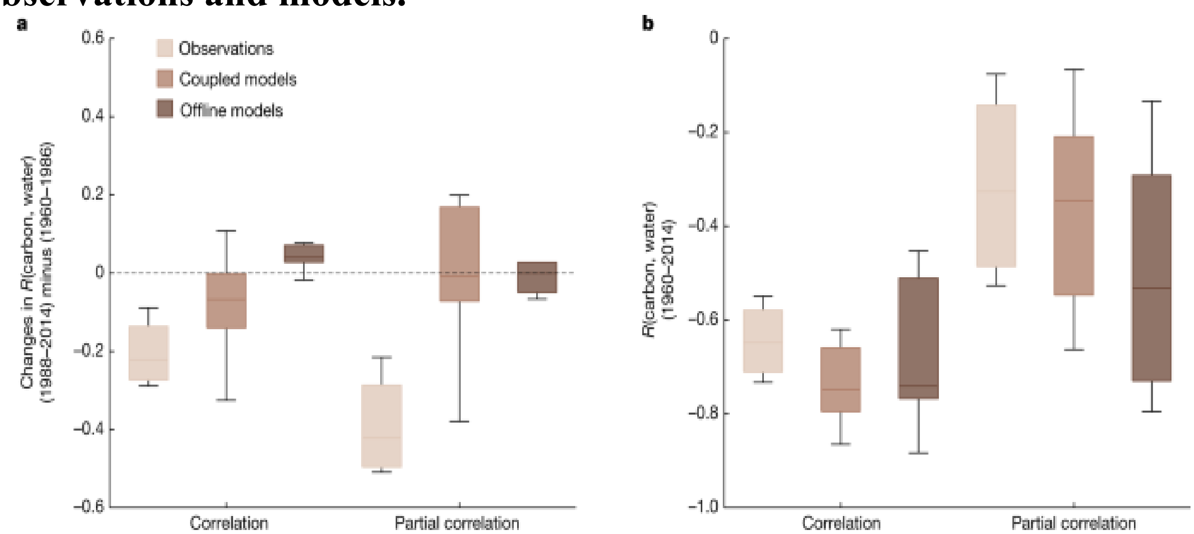
over the region of 2° S–2° N, 110° E–90° W lies in the Eastern Pacific (east of 150° W) and Nino3 index exceeds 1 s.d. Year is considered Central Pacific (CP) ENSO when the corresponding largest DJF SST anomaly lies in the Central Pacific (west of 150° W) and Nino4 index exceeds 1 s.d. Volcano years are excluded from analyses. Grey vertical dashed lines connect the symbols of high spatial coherence and ENSO. **b**, Fraction of years with high spatial coherence within the first 30-yr period (1960–1989) and within the recent 30-yr period (1989–2018). Neutral years are identified as years not in the EP ENSO or CP ENSO state. **c**, Dependence of $R_{\mathrm{WS,CGR}}$ and $R_{\mathrm{WS,CGR|T}}$ on the spatial coherence of WS anomaly. $^{**}P < 0.05$ (significant correlation).

Diagnosis of CMIP6 models

The observational relationships between tropical climate and CGR are valuable metrics to diagnose the ability of models to simulate terrestrial climate–carbon interannual variability. We thus investigate whether state-of-the-art models participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6) can capture this observed feature. We calculate the interannual correlation between tropical WS (using tropical total soil moisture as proxy) and simulated global net ecosystem exchange (NEE, ecosystem respiration minus photosynthesis) estimated by an ensemble of nine coupled Earth System Models (ESMs) and six offline Land Surface Models (LSMs) in the historical 1960–2014 period (Methods). We find that the (partial) correlations between simulated tropical soil moisture and global NEE are persistently high and thus remain almost unchanged over time in most ESMs and LSMs (Fig. 3a and Extended Data Fig. 5). Also, most models do not reproduce the increased sensitivity of NEE to tropical soil moisture, even though they differ largely in the absolute magnitude of the sensitivity (Extended Data Fig. 6). These results indicate that models do not capture the observed emerging enhancement of tropical water–carbon coupling over time, although models roughly capture the sign and strength of this interannual water–carbon relationship during 1960–2014 (Fig. 3b). The ability of models to reproduce observed tropical water–carbon coupling depends not only on the simulation of terrestrial water availability but also on process representations of the response of the carbon cycle to climate. We further find that models might not represent the latter part well because the

modelled water–carbon coupling is stable over time, regardless of the large differences among simulated soil moisture. Specifically, the dominant spatial patterns of simulated soil moisture anomalies are largely divergent among models (Supplementary Figs. 7 and 8), although simulated soil moisture from all offline models can generally capture ENSO teleconnections (Supplementary Fig. 9). For coupled ESMs, the underpinning reason is more complex, for instance, they have known issues in simulating the probability of occurrence of historical EP ENSO and CP ENSO[37]. Compared to ecosystem respiration, the simulated response of ecosystem gross primary production to soil moisture in models is more consistent (Supplementary Figs. 10 and 11). Further including possible modelled carbon fluxes from fire and other disturbances, that is, replacing NEE with net biome production, cannot help to explain the failure of models to reproduce the intensified water–carbon coupling (Supplementary Fig. 12). Models also possibly lack some critical process representations[11], such as the parameterization of deep water uptake[38], tree mortality[39] and plant root adaptation-related characteristics[40]. Therefore, the ability of models to project future terrestrial carbon–climate feedbacks are subject to uncertainties and these results call for improvements on water–carbon interactions to better constrain projections.

Fig. 3: Historical tropical land water–carbon interannual correlations in observations and models.



a, Changes in interannual correlations and partial correlations between tropical water availability and land carbon fluxes from the previous 27 years

(1960–1986) to the recent 27 years (1988–2014). For partial correlations, tropical temperature is controlled. All variables are detrended at yearly scale in each corresponding window. For observations, CGR/RLS, reconstructed tropical WS/tropical LagP and tropical temperature are used for computations ($n = 4$). For models, global NEE, tropical total soil moisture and tropical temperature from each model are used for computations (coupled models, $n = 9$; offline models, $n = 6$). Box plots show the distribution of estimates for all models, solid horizontal lines indicate the median values, boxes cover the interquartile range, and vertical lines reach the 5th and 95th percentiles. **b**, Same as **a** but showing the interannual correlations and partial correlations between tropical water availability and land carbon fluxes during 1960–2014.

Direct observations of tropical net and constituent carbon fluxes covering such a long time period are lacking, limiting spatially explicit attributions of changes in carbon fluxes to specific regions and underpinning drivers. Recent evidence shows that aboveground carbon (AGC) fluxes in tropical semi-arid biomes are strongly associated with CGR IAV during 2011–2017[41]. Tropical AGC dynamics are retrieved from microwave satellite observations of vegetation optical depth (VOD). However, new evidence raises a caution about interpreting IAV of VOD-derived AGC (AGC_{VOD}) as biomass IAV alone because this might be more directly linked to soil moisture[42]. Therefore, although we find increased semi-arid CGR-AGC_{VOD} coupling in the recent 15 years (2002–2016) relative to the first 15 years (1989–2003) using the longest available Ku-band VOD following the approach of ref. [41] (Methods and Supplementary Fig. 13), the underpinning interpretations need further validations. Nonetheless, independent analyses from VOD imply that water-sensitive semi-arid ecosystems might have become more important for CGR IAV during the past three decades. Variations in other constituent carbon fluxes are still not available; for instance, the temporal dynamics of large-scale soil respiration and the temperature sensitivity of soil respiration are constrained by data availability and thus remain uncertain[43]. Other possible mechanisms that are not investigated here include synergy effects of other drought-induced disturbances (such as fire effects[44]), lasting effects of tree mortality on carbon uptake[45] and decreasing temperature sensitivity of both tropical photosynthesis and/or soil respiration under global warming[46]. The most

recent 30 years (1989–2018) overlaps with the decadal 'global warming hiatus' (1998–2012) in which natural internal variability, like ENSO, might play a part[47]. The effects of internal variability should not change the increased tropical water control on CGR IAV in the future because we have already taken ENSO into account. Further studies are required to investigate these potential mechanisms, for instance by integrating new observations from flux towers, field experiments and satellites[48] to calibrate process-oriented models.

In summary, we demonstrate that tropical water availability is likely to have increasingly controlled interannual atmospheric CGR over the past 59 years. The dominant climate driver of the interannual variation of terrestrial carbon cycle has already shown a tendency to shift from temperature to water, suggesting rising water limitations on tropical terrestrial carbon sink. We therefore also partly reconcile the debate of water versus temperature controls for the land carbon cycle[7,9,11] from the perspective of the considered time frames, in addition to the recent view that land–atmosphere feedback matters[20]. As the sensitivity of terrestrial carbon uptake to temperature is usually used as the metric to diagnose or constrain the terrestrial carbon–climate feedbacks[49], we call for more attention to the relevance of tropical water in predicting next-year atmospheric CGR and we suggest it is timely to introduce water-based constraints on future tropical terrestrial carbon–climate feedbacks. Uncertainties in terrestrial carbon–climate feedbacks strongly affect the assessment of the magnitude of emission reductions required to achieve any global temperature target. Hence, the failure of state-of-the-art models to capture the observed increasingly negative coupling between tropical land water and interannual CGR calls for a better characterization of relevant processes to improve the representation of the terrestrial carbon cycle in ESMs and climate projections.

Methods

CGR

Annual global atmospheric CGR spanning from 1960 to 2018 is obtained from the Greenhouse Gas Marine Boundary Layer Reference of the National Oceanic and Atmospheric Administration (NOAA/ESRL)[50]. According to the guideline, the annual CGR in a given year is the difference in CO_2 concentration between the end of December and the start of January of that year. In addition, we also use the estimated RLS from the latest Global Carbon Budget 2020[1] to verify the robustness of our main finding (Extended Data Fig. 1b). RLS is inferred as a residual between emissions, atmospheric CO_2 accumulation and the ocean sink.

Climate data

Terrestrial water storage and lagged precipitation

The twin GRACE satellites provide the measurement of changes in terrestrial WS at monthly scale since March 2002[51]. In particular, terrestrial WS is the sum of all above- and below-surface WS, including soil moisture, groundwater, snow, ice and water stored in vegetation and rivers and lakes. To complement the shorter record of observations provided by the GRACE satellites, we use a recently published statistical reconstruction of terrestrial WS that trained on the twin GRACE satellites[18]. The reconstructed terrestrial WS is based on two different GRACE solutions and three different meteorological forcing datasets. Here, we mainly use the ensemble mean of all members. The details of the statistical approach are documented in ref. [18]. Validation against original terrestrial WS supports the reliability of reconstructions in reproducing historical signals at yearly time scale, including tropics (Supplementary Fig. 14 and Supplementary Table 1). The reconstructed terrestrial WS dataset has a spatial resolution of 0.5° × 0.5° and a temporal resolution of one day from 1901 to 2018.

In addition, direct observations of LagP over tropical land are found to capture IAV of aggregated tropical WS anomaly at yearly time scale (Extended Data Fig. 3a and Supplementary Table 1). For instance, the LagP in 2018 is the sum up of the precipitation from July in 2017 to June in 2018. Moreover, consistent with previous findings[7,17], the IAV of tropical land lagged precipitation correlates with CGR well (Extended Data Fig. 3b). This

helps to explain the rationality of this relationship from the process of WS memory. Therefore, LagP is identified as another efficient proxy for aggregated tropical WS anomaly IAV. Precipitation is obtained from station-based Climate Research Unit (CRU) TS4.03 (ref. [52]) and Global Precipitation Climatology Centre (GPCC) Full Data Monthly v.2020. To further confirm that the data quality of station-based tropical precipitation is reliable, we compare station-based CRU precipitation with satellite-based Tropical Rainfall Measuring Mission (TRMM) precipitation (Extended Data Fig. 7a); we also compare CRU precipitation with GPCC precipitation which has much larger gauge stations than CRU (Extended Data Fig. 7b). In fact, compared to the 2000s to 2010s (in which satellite observations confirm the reliability of station-based tropical precipitation IAV), the number of gauge stations was much larger in the 1960s–1990s (Extended Data Fig. 7c). These validations suggest that station-based tropical precipitation IAV is reliable during 1960–2018. TRMM 3B43 precipitation dataset has a spatial resolution of 0.25° × 0.25° and a temporal resolution of 1 month from 1998 to 2019. Both CRU and GPCC have a spatial resolution of 0.5° × 0.5° and a temporal resolution of 1 month from 1891 to 2018.

Therefore, we use yearly reconstructed tropical WS and LagP to indicate the IAV of tropical terrestrial water availability anomaly.

Temperature

Temperature is obtained from CRU TS4.03 (ref. [52]). The Berkeley Earth global surface temperature is also used for a robust test[53]. The two temperature datasets both have a spatial resolution of 0.5° × 0.5° and a temporal resolution of 1 month from 1901 to 2018.

Regional domain definition

Tropical lands are defined as the spatial average over all the vegetated land areas between 24° N and 24° S, as in ref. [7]. Tropical semi-arid region domain consists of shrubland and (woody) savannah, which is identified according to the land-cover classification map from MODIS (MCD12C1, type3). The map was regridded using a majority filter to a spatial resolution of 0.5° × 0.5°.

ENSO indices

The ENSO is Earth's most important source of interannual climate variability. The areal averages of sea surface temperature (SST) anomalies relative to a long-term average climatology are used to characterize ENSO. The time period of 1960–2018 is used as the climatology here. The SST anomalies over Nino3.4 region (5° N–5° S,120° W–170° W) are most commonly used and included the signals of both EP ENSO and CP ENSO[54]. Year is considered EP ENSO when the largest DJF (December to February) SST anomaly over the region of 2° S–2° N, 110° E–90° W lies in the EP (east of 150° W) and Nino3 index exceeds 1 s.d. Year is considered CP ENSO when the corresponding largest DJF SST anomaly lies in the CP (west of 150° W) and Nino4 index exceeds 1 s.d. We note that the identification of ENSO types could vary with the method used[27]. This study uses the Extended Reconstruction Sea Surface Temperature (ERASST v.5) from the National Oceanic and Atmospheric Administration (NOAA). This dataset has a temporal resolution of 1 month from 1855 to the present and a spatial resolution of 2° × 2° (ref. [55]).

CMIP6 models

Coupled ESM

Nine coupled ESMs participating in the sixth phase of Coupled Model Intercomparison Project (CMIP6) are used: CESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MPI-ESM1-2-LR, UKESM1-0-LL, ACCESS-ESM1-5, CanESM5, MIROC-ES2L and NorESM2-LM. Coupled ESMs allow for feedbacks between the physical climate and the biological and chemical processes in the ocean and on land. We adopted data output from the 'historical' scenario (1960–2014) with one ensemble member for each model. Climate and land carbon sinks are simulated with all-forcings, including both the natural causes (for example, volcanic eruptions and solar variability) and human factors (for example, CO_2 concentration, aerosols and land use) over the period 1850–2014. In coupled ESMs, the carbon cycle is coupled to the climate system.

Offline LSM

Six offline LSMs from the Land Surface, Snow and Soil Moisture Model Intercomparison Project (LS3MIP) are used here[56]: CESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MPI-ESM1-2-LR, UKESM1-0-LL and CMCC-ESM2. The offline LSMs account for land-use changes but do not include local land–atmosphere feedbacks. We adopted data output from the 'Land-Hist' scenario (1960–2014) with one ensemble member for each model, for which the atmospheric forcing, vegetation, soil, topography and land/sea mask data were prescribed following the protocol used for the CMIP6 DECK simulations. The atmospheric forcing comes from the Global Soil Wetness Project phase three (GSWP3), which is a dynamically downscaled and bias-corrected version of the Twentieth Century Reanalysis[57]. Spin-up of the land-only simulations follow the protocol of the project 'Trends and drivers of the regional scale sources and sinks of carbon dioxide' (TRENDY)[58].

Following previous efforts[11,59], to enable a fair comparison of the water–carbon relationship between observations and models, we use the sum of soil moisture in all layers and snow water equivalent as modelled terrestrial WS. In the tropics, snow water equivalent is negligible.

Partial correlation

Partial correlation is used here to directly check whether increasingly negative water–carbon coupling is influenced by confounding water–temperature coupling. However, using specific values of $R_{W,CGR|T}$ to conclude the sign and strength of total water impacts on CGR is not suggested. $R_{W,CGR|T}$ isolates water impacts on CGR from confounding water–temperature coupling by linearly removing all temperature-related covariations. However, given the well-documented soil moisture–atmosphere feedbacks[21], temperature variability actually includes many feedbacks from soil moisture (for example, hot extremes at tropical semi-arid regions) and removing all of them would indirectly remove some water impacts on CGR because of their physical connection. In addition, models do not reproduce the intensified water–carbon coupling but roughly capture the sign and strength of long-term tropical water–carbon coupling during

1960–2018, thus providing insights into underpinning processes. Model factorial experiments show that removing soil moisture interannual variability suppresses land carbon uptake variability by about 90%, whereas tropical mean temperature remains unchanged (Extended Data Fig. 10 in ref. [20]). Therefore, ref. [20] suggest that tropical mean temperature might not represent a mechanistic climatic driver for land carbon uptake variability. Hence, $R_{W,CGR|T}$ is an insufficient and less accurate measure to infer the sign and strength of independent water impacts on CGR and underestimate water impacts on CGR in phases in which temperature control is dominant. Nonetheless, their relative changes are useful in this study and support the finding that water–carbon coupling has become increasingly negative in the recent past (1989–2018) compared to previous climate conditions (1960–1989).

Empirical orthogonal function

The method of empirical orthogonal function (EOF) analysis can deconvolve the spatiotemporal variability of a signal into orthogonal modes, each indicated by a principal spatial pattern and the corresponding principal component time series. It is widely used to study spatial patterns of climate variability and how they change with time[60,61]. We perform the EOF analysis on simulated tropical soil moisture from CMIP6 models.

Ridge regression

In the presence of collinearity, using the OLS estimator can lead to regression coefficient estimates that have large sampling variability and even a wrong sign. Ridge regression is a common technique to be used to address the issues arising through collinearity[62]. In ridge regression, a penalty term is added to the loss function to shrink the regression coefficients[63]. The amount of shrinkage is defined by a regularization parameter that was chosen by us in a cross-validation approach. The data were randomly split 25 times into training and validation sets and, for each split, the performance for the validation test set (mean squared error) was assessed for 100 different regularization parameters spaced evenly on a log scale. The best-performing regularization parameters were selected for each split and the average

between them was retained for the final model. To assess the uncertainty in the regression coefficient estimates, we relied on bootstrapping, meaning that we randomly sampled the data 5,000 times and estimated the regression coefficients for each sample.

Spatial coherence

To quantify the degree of spatial coherence of yearly tropical WS anomaly, following ref. [9], we calculate a large covariance matrix of all grid cells versus all grid cells for tropical WS anomaly. Each element in this covariance matrix is termed as $c_{i,j}$ as follows:

$${c}_{i,j}={\rm{cov}}({{\rm{WS}}}_{i},{{\rm{WS}}}_{j})$$
(1)

where i and j indicate the two grid cells that used to calculate covariance, WS_i and WS_j are the corresponding yearly WS anomaly time series in the specified time period. Then, we summed all, positive and negative covariance terms (termed as tcov, $tcov^+$ and $tcov^-$), respectively, as follows:

$$\begin{array}{c}{\rm{tcov}}=\sum _{i=1}\sum _{j\ne i}| {c}_{i,j}| \\ \,{{\rm{tcov}}}^{+}=\sum _{i=1}\sum _{j\ne i}{c}_{i,j}| {c}_{i,j} > 0\\ \,{{\rm{tcov}}}^{-}=\sum _{i=1}\sum _{j\ne i}{c}_{i,j}| {c}_{i,j} < 0\end{array}$$
(2)

The variances in the diagonal of the covariance matrix (where $i=j$) were excluded because they are always positive and do not contribute to the estimate of spatial coherence. Finally, the spatial coherence of WS anomaly was defined as the following equation:

$${\rm{Spatial}}\,{\rm{coherence}}=\frac{{{\rm{tcov}}}^{+}+{{\rm{tcov}}}^{-}}{{\rm{tcov}}}\times 100 \% $$
(3)

In theory, 100% indicates all grid cells covariate in the same sign, that is, highest spatial coherence. Lower values indicate that total positive

covariances are counterbalanced by total negative covariances, that is, lower spatial coherence.

Vegetation optical depth and aboveground carbon

VOD retrieved from microwave satellite observations is linked to the water content of vegetation mass and offers possibilities for monitoring AGC dynamics[41]. We used recently published long-term VOD products from the VOD Climate Archive (VODCA), which combines VOD retrievals that have been derived from multiple sensors (SSM/I, TMI, AMSR-E, WindSat and AMSR2) using the Land Parameter Retrieval Model[64]. For time completeness, we used the longest available VOD estimated from Ku-band which covers the period 1988–2016. To estimate the tropical AGC, following the approach of ref. [65], we first fitted a four-parameter empirical function by calibrating tropical VOD against the tropical AGC benchmark map from ref. [66] in 2000 as follows:

$${\rm{AGC}}=a\times \frac{\arctan (b\times ({\rm{VOD}}-c))-\arctan (-b\times c)}{\arctan (b\times ({\rm{Inf}}-c))-\arctan (-b\times c)}+d$$
(4)

where a, b, c and d are four best-fit parameters and Inf is set to 10^{10}. AGC density ($MgC\ ha^{-1}$) was derived by multiplying the original aboveground biomass density values by a factor of 0.5 (ref. [67]). For Ku-VOD and AGC, data are aggregated to the spatial resolution of $0.5° \times 0.5°$. The spatial scatter plot of VOD and AGC clearly demonstrates the good relationship between VOD and AGC (coefficient of determination $R^2 = 0.76$, $P < 0.01$; Supplementary Fig. 15). It seems the performance of VODCA Ku-VOD is less comparable to L-VOD ($R^2 = 0.81$) (ref. [41]), which was considered to be more sensitive to AGC in high biomass regions. However, L-VOD has only been available since 2010 and thus is not used here. Finally, we apply this empirical function to convert VOD to AGC from 1988 to 2016.

Data availability

All the datasets used here are publicly available. Atmospheric CO_2 observations are available at https://gml.noaa.gov/ccgg/; GRACE observations of terrestrial WS are available at https://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/; GRACE-REC terrestrial WS are available at https://doi.org/10.6084/m9.figshare.7670849; CRU climate datasets are available at https://www.uea.ac.uk/groups-and-centres/climatic-research-unit; GPCC precipitation dataset is available at https://www.dwd.de/EN/ourservices/gpcc/gpcc.html; TRMM precipitation dataset is available at https://disc.gsfc.nasa.gov/datasets/; Berkeley Earth climate datasets are available at http://berkeleyearth.org/; ERASST v.5 are available at https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html; CMIP6 model outputs are available at https://pcmdi.llnl.gov/CMIP6/; and VODCA products are available at https://doi.org/10.5281/zenodo.2575599.

Code availability

Codes are available through Zenodo at https://doi.org/10.5281/zenodo.6447779.

References

1. Friedlingstein, P. et al. Global carbon budget 2020. *Earth Syst. Sci. Data* **12**, 3269–3340 (2020).

2. Friedlingstein, P. et al. Climate–carbon cycle feedback analysis: results from the (CMIP)-M-4 model intercomparison. *J. Clim.* **19**, 3337–3353 (2006).

3. Bousquet, P. et al. Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science* **290**, 1342–1346 (2000).

4. Piao, S. L. et al. Interannual variation of terrestrial carbon cycle: Issues and perspectives. *Glob. Change Biol.* **26**, 300–318 (2020).

5. Cox, P. M. et al. Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature* **494**, 341–344 (2013).

6. Canadell, J. G. et al. in *Climate Change 2021: The Physical Science Basis* (eds Masson-Delmotte, V. et al.) 673–816 (Cambridge Univ. Press, 2021).

7. Wang, W. L. et al. Variations in atmospheric CO_2 growth rates coupled with tropical temperature. *Proc. Natl Acad. Sci. USA* **110**, 13061–13066 (2013).

8. Anderegg, W. R. L. et al. Tropical nighttime warming as a dominant driver of variability in the terrestrial carbon sink. *Proc. Natl Acad. Sci. USA* **112**, 15591–15596 (2015).

9. Jung, M. et al. Compensatory water effects link yearly global land CO_2 sink changes to temperature. *Nature* **541**, 516–520 (2017).

10. Braswell, B. H., Schimel, D. S., Linder, E. & Moore, B. The response of global terrestrial ecosystems to interannual temperature variability. *Science* **278**, 870–872 (1997).

11. Humphrey, V. et al. Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage. *Nature* **560**, 628-631 (2018).

12. Bacastow, R. B. Modulation of atmospheric carbon dioxide by the Southern Oscillation. *Nature* **261**, 116–118 (1976).

13. Keeling, C. D., Whorf, T. P., Wahlen, M. & van der Plichtt, J. Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature* **375**, 666–670 (1995).

14. Zeng, X. B., Eyre, J. E. J. R., Dixon, R. D. & Arevalo, J. Quantifying the occurrence of record hot years through normalized warming trends. *Geophys. Res. Lett*. **48**, e2020GL091626 (2021).

15. Adams, J. M. & Piovesan, G. Long series relationships between global interannual CO_2 increment and climate: evidence for stability and change in role of the tropical and boreal-temperate zones. *Chemosphere* **59**, 1595–1612 (2005).

16. Wang, X. H. et al. A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature* **506**, 212–215 (2014).

17. Wang, J., Zeng, N. & Wang, M. R. Interannual variability of the atmospheric CO_2 growth rate: roles of precipitation and temperature. *Biogeosciences* **13**, 2339–2352 (2016).

18. Humphrey, V. & Gudmundsson, L. GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. *Earth Syst. Sci. Data* **11**, 1153–1170 (2019).

19. Mercado, L. M. et al. Impact of changes in diffuse radiation on the global land carbon sink. *Nature* **458**, 1014–1087 (2009).

20. Humphrey, V. et al. Soil moisture–atmosphere feedback dominates land carbon uptake variability. *Nature* **592**, 65–69 (2021).

21. Seneviratne, S. I. et al. Investigating soil moisture–climate interactions in a changing climate: a review. *Earth Sci. Rev.* **99**, 125–161 (2010).

22. Luo, X. Z. & Keenan, T. F. Tropical extreme droughts drive long-term increase in atmospheric CO_2 growth rate variability. *Nat. Commun.* **13**, 1193 (2022).

23. Gu, G. J. & Adler, R. F. Precipitation and temperature variations on the interannual time scale: assessing the impact of ENSO and volcanic eruptions. *J. Clim.* **24**, 2258–2270 (2011).

24. Cai, W. J. et al. Changing El Nino-Southern Oscillation in a warming climate. *Nat. Rev. Earth Environ.* **2**, 628–644 (2021).

25. Freund, M. B. et al. Higher frequency of Central Pacific El Nino events in recent decades relative to past centuries. *Nat. Geosci.* **12**, 450–455 (2019).

26. Johnson, N. C. How many ENSO flavors can we distinguish? *J. Clim.* **26**, 4816–4827 (2013).

27. Capotondi, A. et al. Understanding ENSO diversity. *Bull. Am. Meteorol. Soc.* **96**, 921–938 (2015).

28. Ashok, K., Behera, S. K., Rao, S. A., Weng, H. Y. & Yamagata, T. El Nino Modoki and its possible teleconnection. *J. Geophys. Res. Oceans* https://doi.org/10.1029/2006JC003798 (2007).

29. Wang, G. & Hendon, H. H. Sensitivity of Australian rainfall to inter-El Nino variations. *J. Clim.* **20**, 4211–4226 (2007).

30. de Linage, C., Kim, H., Famiglietti, J. S. & Yu, J. Y. Impact of Pacific and Atlantic sea surface temperatures on interannual and decadal variations of GRACE land water storage in tropical South America. *J. Geophys. Res. Atmos* **118**, 10811–10829 (2013).

31. Rodell, M. et al. Emerging trends in global freshwater availability. *Nature* **557**, 651–659 (2018).

32. Wang, J. et al. Enhanced India–Africa carbon uptake and Asia-Pacific carbon release associated with the 2019 extreme positive Indian Ocean dipole. *Geophys. Res. Lett.* **49**, e2022GL100950 (2022).

33. Wang, J. et al. Modulation of land photosynthesis by the Indian Ocean dipole: satellite-based observations and CMIP6 future projections. *Earths Future* **9**, e2020EF001942 (2021).

34. Yeh, S. W. et al. El Nino in a changing climate. *Nature* **461**, 511–570 (2009).

35. Chylek, P., Tans, P., Christy, J. & Dubey, M. K. The carbon cycle response to two El Nino types: an observational study. *Environ. Res. Lett.* **13**, 024001 (2018).

36. Wang, J. et al. Contrasting interannual atmospheric CO_2 variabilities and their terrestrial mechanisms for two types of El Ninos. *Atmos. Chem. Phys.* **18**, 10333–10345 (2018).

37. Timmermann, A. et al. El Nino-Southern Oscillation complexity. *Nature* **559**, 535–545 (2018).

38. Fan, Y., Miguez-Macho, G., Jobbagy, E. G., Jackson, R. B. & Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proc. Natl Acad. Sci. USA* **114**, 10572–10577 (2017).

39. Yang, Y. et al. Post-drought decline of the Amazon carbon sink. *Nat. Commun.* **9**, 3172 (2018).

40. Feddes, R. A. et al. Modeling root water uptake in hydrological and climate models. *Bull. Am. Meteorol. Soc.* **82**, 2797–2809 (2001).

41. Fan, L. et al. Satellite-observed pantropical carbon dynamics. *Nat. Plants* **5**, 944–951 (2019).

42. Konings, A. G., Holtzman, N. M., Rao, K., Xu, L. & Saatchi, S. S. Interannual variations of vegetation optical depth are due to both water stress and biomass changes. *Geophys. Res. Lett.* **48**, e2021GL095267 (2021).

43. Bond-Lamberty, B. New techniques and data for understanding the global soil respiration flux. *Earths Future* **6**, 1176–1180 (2018).

44. Liu, J. J. et al. Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Nino. *Science* https://doi.org/10.1126/science.aam5690 (2017).

45. da Costa, A. C. L. et al. Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytol.* **187**, 579–591 (2010).

46. Huang, M. T. et al. Air temperature optima of vegetation productivity across global biomes. *Nat. Ecol. Evol.* **3**, 772–779 (2019).

47. Medhaug, I., Stolpe, M. B., Fischer, E. M. & Knutti, R. Reconciling controversies about the 'global warming hiatus'. *Nature* **545**, 41–47 (2017).

48. Liu, L. B. et al. Soil moisture dominates dryness stress on ecosystem production globally. *Nat. Commun.* **11**, 4892 (2020).

49. Arora, V. K. et al. Carbon-concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences* **17**, 4173–4222 (2020).

50. Masarie, K. A. & Tans, P. P. Extension and integration of atmospheric carbon-dioxide data into a globally consistent measurement record. *J. Geophys. Res. Atmos.* **100**, 11593–11610 (1995).

51. Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. & Watkins, M. M. GRACE measurements of mass variability in the Earth system. *Science* **305**, 503–505 (2004).

52. Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **7**, 109 (2020).

53. Rohde, R. A. & Hausfather, Z. The Berkeley Earth Land/Ocean temperature record. *Earth Syst. Sci. Data* **12**, 3469–3479 (2020).

54. Trenberth, K. E. The definition of El Nino. *Bull. Am. Meteorol. Soc.* **78**, 2771–2777 (1997).

55. Huang, B. Y. et al. Extended reconstructed sea surface temperature, Version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).

56. van den Hurk, B. et al. LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project—aims, setup and expected outcome. *Geosci. Model Dev.* **9**, 2809–2832 (2016).

57. Compo, G. P. et al. The twentieth century reanalysis project. *Q. J. R. Meteor. Soc.* **137**, 1–28 (2011).

58. Sitch, S. et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015).

59. Wu, R. J., Lo, M. H. & Scanlon, B. R. The annual cycle of terrestrial water storage anomalies in CMIP6 models evaluated against GRACE

data. *J. Clim.* **34**, 8205–8217 (2021).

60. Björnsson, H. & Venegas, S. J. C. R. A manual for EOF and SVD analyses of climatic data. *CCGCR Rep.* **97**, 112–134 (1997).

61. Cai, W. J. et al. ENSO and greenhouse warming. *Nat. Clim. Change* **5**, 849–859 (2015).

62. Dormann, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).

63. Hoerl, A. E. & Kennard, R. W. Ridge regression—biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).

64. Moesinger, L. et al. The global long-term microwave Vegetation Optical Depth Climate Archive (VODCA). *Earth Syst. Sci. Data* **12**, 177–196 (2020).

65. Liu, Y. Y. et al. Recent reversal in loss of global terrestrial biomass. *Nat. Clim. Change* **5**, 470–474 (2015).

66. Avitabile, V. et al. An integrated pan-tropical biomass map using multiple reference datasets. *Glob. Change Biol.* **22**, 1406–1420 (2016).

67. Baccini, A. et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Change* **2**, 182–185 (2012).

Acknowledgements

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access and the multiple funding agencies who support CMIP6 and ESGF. We thank all contributors to the LS3MIP and LMIP experiments. We thank the Global Monitoring Division of

NOAA/Earth System Research Laboratory for providing the atmospheric CO_2 measurements. We thank U. Beyerle, L. Brunner, R. Lorenz and M. Hauser for downloading and processing the CMIP6 and LS3MIP data. We acknowledge the NOAA/OAR.ESRL PSL for providing the ERSST_v5 dataset. We thank M. Hirschi for downloading the VOD data. We acknowledge V. Humphrey and Q. Sun for useful discussions. L.L., R.S.P., P.C., P.F. and S.I.S. acknowledge support from the European Union's Horizon 2020 Research and Innovation Programme (grant no. 821003 (4C)). L.L., P.C. and S.I.S. also acknowledge support from HORIZON.2.5 (grant no. 101056939) (RESCUE).

Funding

Open access funding provided by Swiss Federal Institute of Technology Zurich.

Author information

Authors and Affiliations

1. Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

   Laibao Liu, Ryan S. Padrón, Jonas Schwaab, Lukas Gudmundsson & Sonia I. Seneviratne

2. Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Université Paris Saclay, Gif-sur-Yvette, France

   Philippe Ciais

3. Joint Institute for Regional Earth System Science and Engineering (JIFRESSE), University of California, Los Angeles, Los Angeles, CA, USA

   Mengxi Wu

4. College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

   Pierre Friedlingstein

Contributions

L.L. conceived the original idea. L.L., S.I.S. and P.C. designed the experiments. L.L., S.I.S., P.C., W.M., R.S.P., P.F., J.S. and L.G. performed the research. J.S. carried out ridge regression analyses. L.L. carried out all other analyses. L.L. wrote the paper with contributions from all co-authors.

Corresponding author

Correspondence to Laibao Liu.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Russell Scott, Jun Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Robustness of tropical land climate–carbon interannual correlations.

a, Histograms of climate–carbon interannual correlations in the first three decades (1960–1989) and in the recent three decades (1989–2018), derived using 5000 bootstrapping repeats. Same as Fig. 1c, but tropical water refers to 6-month lagged precipitation from GPCC and tropical temperature is derived from Berkeley Earth global surface temperature. **b**, Same as Fig. 1c, but CGR is replaced with residual land sink (RLS).

Extended Data Fig. 2 Interannual sensitivity of CGR to tropical temperature variations.

a, Interannual sensitivity of CGR to tropical temperature is estimated from the bivariate regression with both tropical temperature and tropical water as predictors. Pre indicates concurrent precipitation. To keep the consistency in method with the previous relevant studies, we first employ the OLS regression. Each dot indicates a 25-year period. The central year of the time window is labelled on the horizontal axis. Shaded areas represent the 95% confidence interval, derived using 5000 bootstrapping repeats. For a better readability, only the 95% confidence interval of temperature sensitivity from the first and third regression function are plotted. **b**, same as a, but using the Ridge regression.

Extended Data Fig. 3 Suitability of using LagP as the poxy of tropical terrestrial water IAV.

a, **b**, IAV of tropical LagP and (a) tropical WS and (b) CGR during 1960–2018. Vertical grey shades indicate three volcanic eruptions (Mount Agung, El Chichón and Pinatubo).

Extended Data Fig. 4 Spatial pattern of WS anomaly at EP ENSO and CP ENSO years.

Year is considered EP ENSO when the largest DJF SST anomaly over the region of 2°S–2°N, 110°E–90°W lies in the Eastern Pacific (Eastern of 150°W) and Nino3 index exceeds one standard deviation. Year is considered CP ENSO when the corresponding largest DJF SST anomaly lies in the Central Pacific (Western of 150°W) and Nino4 index exceeds one standard deviation. Volcano years are excluded from analyses.

Extended Data Fig. 5 Interannual correlation of land carbon fluxes to tropical water in models.

Years labelled on the horizontal axis indicate the central year of the 27-year moving time window (all variables detrended at yearly scale in each corresponding window). Models are based on tropical total soil moisture, tropical temperature and global net ecosystem exchange.

Extended Data Fig. 6 Interannual sensitivity of land carbon fluxes to tropical water in models.

For models, global net ecosystem exchange, tropical total soil moisture, and tropical temperature from each model are used for computations. Univariate and bivariate sensitivity are estimated using the OLS regression and Ridge regression, respectively. The best estimate of sensitivity of CGR to tropical water is shown. ** indicates a significant sensitivity at $P < 0.05$.

Extended Data Fig. 7 Robustness of tropical land LagP IAV.

a, **b**, IAV of tropical LagP in CRU, (a) TRMM during 1999–2018 and (b) GPCC during 1960–2018. **c**, Temporal coverage of the number of gauge stations in GPCC and CRU.

Extended Data Table 1 Interannual correlation and sensitivity of CGR to tropical water

Supplementary information

Supplementary Information Supplementary Figs. 1–15 and Table 1.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Liu, L., Ciais, P., Wu, M. *et al.* Increasingly negative tropical water–interannual CO_2 growth rate coupling. *Nature* **618**, 755–760 (2023). https://doi.org/10.1038/s41586-023-06056-x

- Received: 05 January 2022
- Accepted: 05 April 2023
- Published: 31 May 2023

- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06056-x

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06056-x

- Article
- Open Access
- Published: 17 May 2023

Polygenic scoring accuracy varies across the genetic ancestry continuum

- Yi Ding ,
- Kangcheng Hou,
- Ziqi Xu,
- Aditya Pimplaskar,
- Ella Petter,
- Kristin Boulier,
- Florian Privé,
- Bjarni J. Vilhjálmsson,
- Loes M. Olde Loohuis &
- …
- Bogdan Pasaniuc

Show authors

Nature **volume 618**, pages 774–781 (2023)

- 12k Accesses
- 210 Altmetric
- Metrics details

Abstract

Polygenic scores (PGSs) have limited portability across different groupings of individuals (for example, by genetic ancestries and/or social determinants of health), preventing their equitable use[1,2,3]. PGS portability has typically been assessed using a single aggregate population-level statistic (for example, R^2)[4], ignoring inter-individual variation within the population. Here, using a large and diverse Los Angeles biobank[5] (ATLAS, $n = 36{,}778$) along with the UK Biobank[6] (UKBB, $n = 487{,}409$), we show that PGS accuracy decreases individual-to-individual along the continuum of genetic ancestries[7] in all considered populations, even within traditionally labelled 'homogeneous' genetic ancestries. The decreasing trend is well captured by a continuous measure of genetic distance (GD) from the PGS training data: Pearson correlation of −0.95 between GD and PGS accuracy averaged across 84 traits. When applying PGS models trained on individuals labelled as white British in the UKBB to individuals with European ancestries in ATLAS, individuals in the furthest GD decile have 14% lower accuracy relative to the closest decile; notably, the closest GD decile of individuals with Hispanic Latino American ancestries show similar PGS performance to the furthest GD decile of individuals with European ancestries. GD is significantly correlated with PGS estimates themselves for 82 of 84 traits, further emphasizing the importance of incorporating the continuum of genetic ancestries in PGS interpretation. Our results highlight the need to move away from discrete genetic ancestry clusters towards the continuum of genetic ancestries when considering PGSs.

Main

PGSs—estimates of an individual's genetic predisposition for complex traits and diseases (that is, genetic liability; also referred to as genetic value)—have garnered tremendous attention recently across a wide range of fields, from personalized genomic medicine[4,8,9,10] to disease risk prediction and prevention[11,12,13,14] to socio-genomics[3,15]. However, the variation in PGS performance across different genetic ancestries and/or socio-demographic features (for example, sex, age and social determinants of health)[2] poses a critical equity barrier that has prevented widespread adoption of PGSs. Similar portability issues have also been reported for non-genetic clinical models[16,17,18]. The interpretation and application of PGSs are further

complicated by the conflation of genetic ancestries with social constructs such as nationality, race and/or ethnicity. Here we investigate PGS performance across genetically inferred ancestry (GIA), which describes the genetic similarity of an individual to a reference dataset (for example, 1000 Genomes[19]) as inferred by methods such as principal component analysis (PCA); GIAs do not represent the full genetic diversity of human populations.

Genetic prediction and its accuracy (or reliability) have been extensively studied in agricultural settings with a focus on breeding programmes[20,21,22,23]. At the population level, PGS accuracy can be expressed as a function of heritability, training sample size and the number of markers used in the predictor in single[24,25,26] or multi-population settings with or without effect size heterogeneity[27]. At the individual level, accuracy of genetic prediction from pedigree data[28,29,30] can be derived as a function of the inverse of the coefficient matrix of mixed-models equations, whereas accuracy of genetic prediction using whole-genome genetic data can be derived similarly, with the pedigree matrix replaced with the genomic relationships matrix[21,22,23,27,31,32] among training and testing individuals. Simulations guided by dairy breeding programmes showcase that genomic prediction accuracy varies with genetic relatedness of the testing individual to the training data[33,34] as well as across generations, owing to the decay of genetic relationships[35].

In humans, PGS performance evaluation has traditionally relied on population-level accuracy metrics (for example, R^2)[2,4]. PGS accuracy decays as the target populations become more dissimilar from the training data using either relatedness[36,37] or continental or subcontinental ancestry groupings[1,38,39,40]; the decay may be explained by differences in linkage disequilibrium, minor allele frequencies and/or heterogeneity in genetic effects due to gene–gene and gene–environment interactions[41]. However, population-level metrics of accuracy provide only an aggregate (average) metric for all individuals in the population, thus implicitly assuming some level of homogeneity across individuals[2,4,42]. Homogeneous populations are an idealized concept that only roughly approximate human data; human diversity exists along a genetic ancestry continuum without clearly defined clusters and with various correlations between genetic and socio-

environmental factors[7,42,43,44,45,46]. Grouping individuals into discrete GIA clusters obscures the impact of individual variation on PGS accuracy. This is evident among individuals with recently admixed genomes for which genetic ancestries vary individual-to-individual and locus-to-locus in the genome. For example, a single population-level PGS accuracy estimated across all African Americans overestimates PGS accuracy for African Americans with large proportions of African GIA[40]; likewise, coronary artery disease PGS performs poorly in Hispanic individuals with high proportions of African GIA[47]. The genetic ancestry continuum affects PGS accuracy even in traditionally labelled ‘homogeneous’ or ‘non-admixed’ populations. For example, PGS accuracy decays across a gradient of subcontinental ancestries within Europe as the target cohorts become more genetically dissimilar from the PGS training data[39,45]. Assessing PGS accuracy using population-level metrics is further complicated by technical issues in assigning individuals to discrete clusters of GIA. Different algorithms and/or reference panels may assign the same individual to different clusters[39,42,48], leading to different PGS accuracies. Moreover, many individuals are not assigned to any cluster owing to limited reference panels used for genetic ancestry inference[5,39], leaving such individuals outside PGS characterization. This poses equity concerns as it limits PGS applications only to individuals within well-defined GIAs.

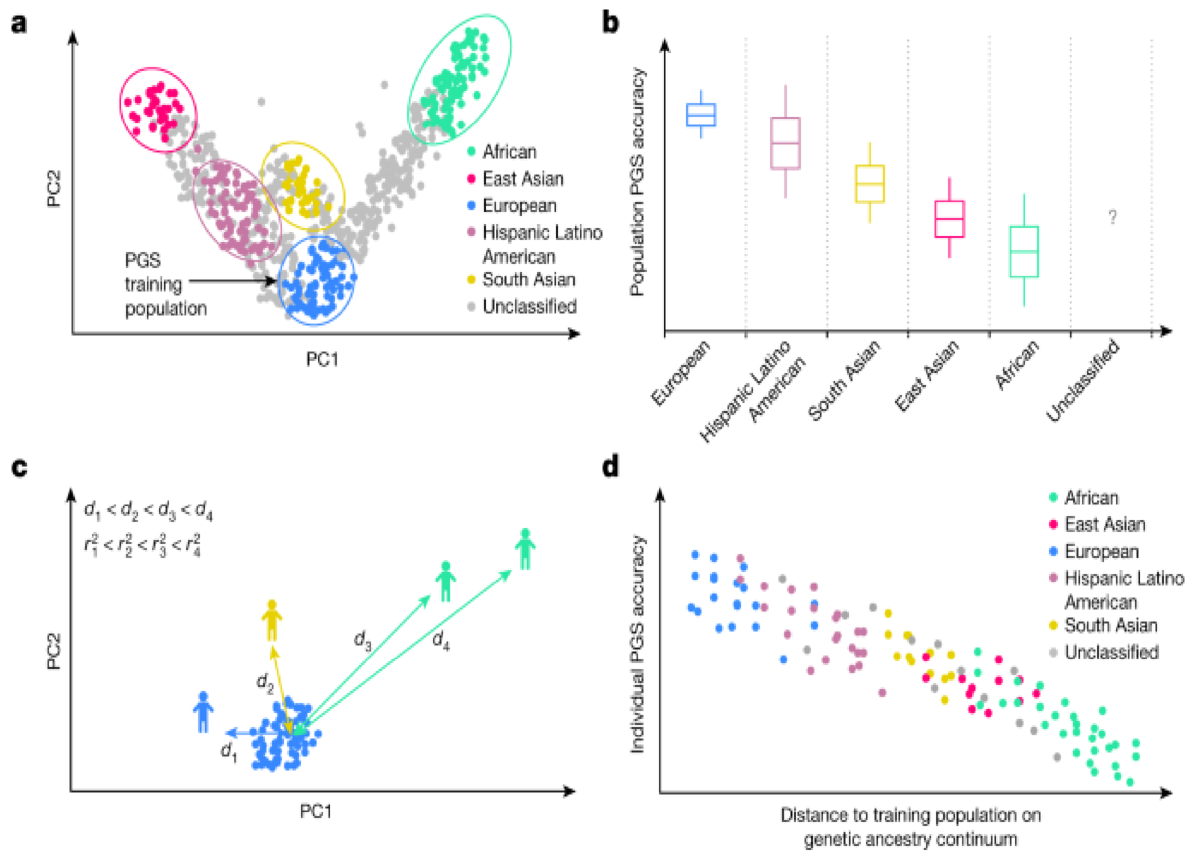
Here we leverage classical theory[28,29,30] and methods that characterize PGS performance at the level of a single target individual[49] to evaluate the impact of the genetic ancestry continuum on PGS accuracy. We use simulations and real-data analyses to show that PGS accuracy decays continuously individual-to-individual across the genetic continuum as a function of GD from the PGS training data; GD is defined as a PCA projection of the target individual on the training data used to estimate the PGS weights. We leverage a large and diverse Los Angeles biobank at the University of California, Los Angeles[5] (ATLAS, $n = 36{,}778$) along with the UK Biobank[6] (UKBB, $n = 487{,}409$) to investigate the interplay between genetic ancestries and PGS for 84 complex traits and diseases. The accuracy of PGS models trained on individuals labelled as white British (WB; see Methods for naming convention used in this work) in the UKBB ($n = 371{,}018$) is negatively correlated with GD for all considered traits (average Pearson $R = -0.95$ across 84 traits), demonstrating pervasive individual variation in PGS

accuracy. The negative correlation remains significant even when restricted to traditionally defined GIA clusters (ranging from $R = -0.43$ for East Asian GIA to $R = -0.85$ for the African American GIA in ATLAS). On average across the 84 traits, when rank-ordering individuals according to distance from training data, PGS accuracy decreases by 14% in the furthest versus closest decile in the European GIA. Notably, the furthest decile of individuals of European ancestries showed similar accuracy to the closest decile of Hispanic Latino individuals. Characterizing PGS accuracy across the continuum allows the inclusion of individuals unassigned to any GIA (6% of all ATLAS), thus allowing more individuals to be included in PGS applications. Finally, we explore the relationship between GD and PGS estimates themselves. Of 84 PGSs, 82 show significant correlation between GD and PGS with 30 showing opposite correlation (GD, trait) versus (GD, PGS); we exemplify the importance of incorporating GD in interpretation of PGSs using height and neutrophils in the ATLAS data. Our results demonstrate the need to incorporate the genetic ancestry continuum in assessing PGS performance and/or bias.

Overview of the study

PGS accuracy has conventionally been assessed at the level of discrete GIA clusters using population-level metrics of accuracy. Individuals from diverse genetic backgrounds are routinely grouped into discrete GIA clusters using computational inference methods such as PCA[50] and/or admixture analysis[51] (Fig. 1a). Population-level metrics of PGS accuracy are then estimated for each GIA cluster and generalized to everyone in the cluster (Fig. 1b). This approach has three major limitations: the inter-individual variability within each cluster is ignored; the GIA cluster boundary is sensitive to algorithms and reference panels used for clustering; and a substantial proportion of individuals may not be assigned to any GIA owing to a lack of reference panels for genetic ancestry inference (for example, individuals of uncommon or admixed ancestries).

Fig. 1: Illustration of population-level versus individual-level PGS accuracy.



a, Discrete labelling of GIA with PCA-based clustering. Each dot represents an individual. The circles represent arbitrary boundaries imposed on the genetic ancestry continuum to divide individuals into different GIA clusters. The colour represents the GIA cluster label. The grey dots are individuals who are left unclassified. **b**, Schematic illustrating the variation of population-level PGS accuracy across clusters. The box plot represents the PGS accuracy (for example, R^2) measured at the population level. The question mark emphasizes that the PGS accuracy for unclassified individuals is unknown owing to the lack of a reference group. Grey dashed lines emphasize the categorical nature of GIA clustering. **c**, Continuous labelling of everyone's unique position on the genetic ancestry continuum with a PCA-based GD. The GD is defined as the Euclidean distance of an individual's genotype from the centre of the training data when projected on the PC space of training genotype data. Everyone has their own unique GD, ${d}_{i}$, and individual PGS accuracy, ${r}_{i}^{2}$. **d**, Individual-level PGS accuracy decays along the genetic ancestry continuum. Each dot represents an individual and its colour represents the assigned GIA label. Individuals labelled with the same ancestry spread out on the genetic

ancestry continuum, and there are no clear boundaries between GIA clusters. This figure is illustrative and does not involve any real or simulated data.

Here we evaluate PGS accuracy across the genetic ancestry continuum at the level of a single target individual. We model the phenotype of individual i as ${y}_{i}={x}_{i}^{{\rm{\top }}}\beta +{{\epsilon }}_{i}$, in which ${x}_{i}$ is an $M\times 1$ vector of standardized genotypes for M variants, β is an $M\times 1$ vector of standardized causal effects, and ${{\epsilon }}_{i}$ is random noise. Under a random effects model, genetic liability ${g}_{i}\,=$ ${x}_{i}^{{\rm{\top }}}\beta$ and its PGS estimate ${\hat{g}}_{i}=E({x}_{i}^{{\rm{\top }}}\beta |D)$ are random variables for which the randomness comes from β and training data D ($D=({X}_{{\rm{train}}},{y}_{{\rm{train}}}\left.\right))$). We define the individual PGS accuracy as the correlation of an individual's genetic liability and PGS estimate with the following equation in consistence with classical theory[28,32,52]:

$${r}_{i}^{2}(\,{g}_{i},{\hat{g}}_{i})=\frac{{{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}{({g}_{i},{\hat{g}}_{i})}^{2}}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }(\,{g}_{i}){{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}(\,{\hat{g}}_{i})}=1-\frac{{E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)}$$
(1)

Under an infinitesimal assumption for which all variants are causal and drawn from a normal distribution $N(0,{{\sigma }}_{\beta }^{2})$, the analytical form of PGS accuracy can be derived as:

$${r}_{i}^{2}(\,{g}_{i},{\hat{g}}_{i})=1-\frac{{{\sigma }}_{e}^{2}{\sum }_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}}{{{\sigma }}_{\beta }^{2}{x}_{i}^{{\rm{\top }}}{x}_{i}}=1-\frac{{{\sigma }}_{e}^{2}}{{{\sigma }}_{\beta }^{2}}\frac{{\sum }_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}}{{x}_{i}^{{\rm{\top }}}{x}_{i}}$$

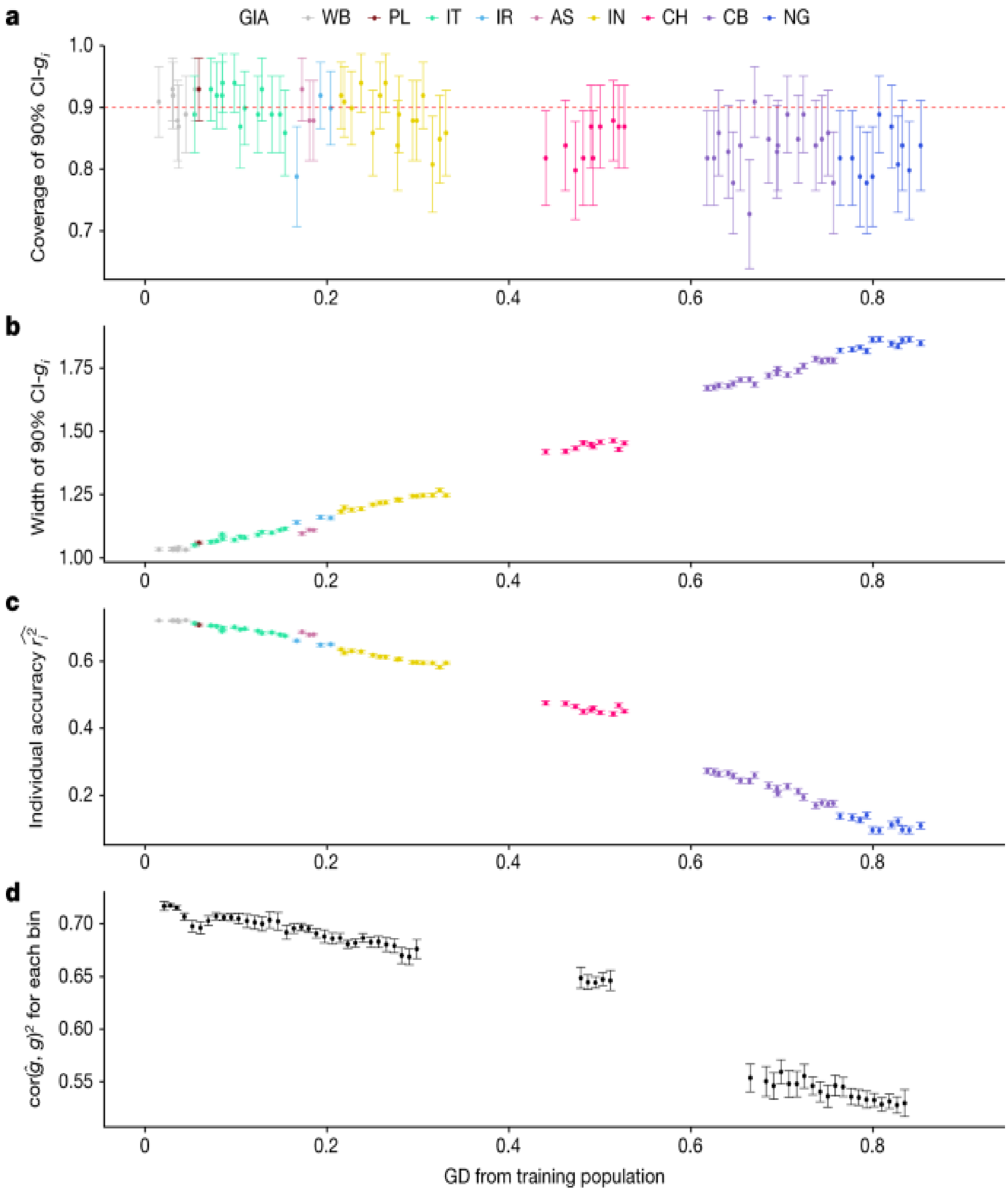
in which ${\sigma }_{\beta }^{2}$ is per single nucleotide polymorphism (SNP) heritability; ${\sigma }_{{\rm{e}}}^{2}$ is the variance of residual environmental noise; ${v}_{j}$ and ${\lambda }_{j}$ are the *j*th eigenvector and eigenvalues of training genotype data and *J* is the total number of eigenvectors. ${\sum }_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}$ is the squared Mahalanobis distance of the testing individual i from the centre of the training genotype data on its principal component (PC) space; and ${x}_{i}^{{\rm{\top }}}{x}_{i}$ is the sum of squared genotypes across all variants. Empirically, the ratio of the squared Mahalanobis distance to the sum of squared genotypes is highly correlated with the Euclidean distance of the individual from the training data on that PC space ($R = 1$, $P < 2.2 \times 10^{-16}$ in the UKBB). Given that this metric of accuracy is highly dependent on the GD from the training data, we term it the panel distance ${r}_{i}^{2}$. In practice, we use LDpred2 to estimate ${E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))$ (refs. [49,53]) and approximate ${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ as the heritability of the phenotype[30] (Methods). As a continuous GD, we use ${d}_{i}=\sqrt{{\sum }_{j=1}^{J}{({x}_{i}^{{\rm{\top }}}{v}_{j})}^{2}}$ with *J* set to 20 (Fig. 1c,d and Methods). We note two caveats of individual PGS accuracy: first, the genetic effects are assumed to be the same for all individuals regardless of their genetic ancestry background; second, the SNPs used for PGS training may not fully capture trait heritability. Therefore, the metric we proposed here is an upper bound of genetic prediction accuracy (Supplementary Note).

PGS performance is calibrated in simulations

First, we evaluated calibration of the posterior variance of genetic liability ${E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))$ estimated by LDpred2 for individuals at various GDs from the UKBB WB training data by checking the calibration of the 90% credible intervals (Fig. 2a). We simulated 100 phenotypes at heritability ${h}_{g}^{2}=0.25$ and proportion of causal variants ${p}_{{\rm{causal}}}=1 \%$ for all individuals in the UKBB, assuming shared causal variants and homogeneous causal effect sizes for individuals from various genetic backgrounds (see Methods). Overall, the 90% credible

intervals are approximately well calibrated (that is, the 90% credible interval overlaps with the true genetic liability across 90 of 100 replicates, for all individuals, regardless of their GD from the training population or GIA labels; Fig. 2a). For example, when individuals are binned into 10 deciles based on their GD from the training population, the average empirical coverage of the 90% credible intervals is 89.7% (s.d. 2.6%) for individuals from the closest decile (composed of 96.9% individuals labelled as WB, 3.1% labelled as PL under a discrete view of ancestries; see detailed naming convention in Methods) compared to the average empirical coverage of 82.4% (s.d. 4.6%) for individuals from the furthest decile (composed of 19.9% individuals labelled as CB and 80.1% labelled as NG).

Fig. 2: PGS performance is calibrated across GD in simulations using UKBB data.



a, The 90% credible intervals of genetic liability (CI-g_i) are well calibrated for testing individuals at all GDs. The red dashed line represents the expected coverage of the 90% CI-g_i. Each dot represents a randomly selected UKBB testing individual. For each dot, the *x*-axis is its GD from the training data, the *y*-axis is the empirical coverage of the 90% CI-g_i

calculated as the proportion of simulation replicates for which the 90% CI-g_i contain the individual's true genetic liability, and the error bars represent the mean ±1.96 standard error of the mean (s.e.m.) of the empirical coverage calculated from 100 simulations. **b**, The width of the 90% CI-g_i increases with GD. For each dot, the *y*-axis is the average width of the 90% CI-g_i across 100 simulation replicates, and the error bars represent ±1.96 s.e.m. **c**, Individual PGS accuracy decreases with GD. For each dot, the *y*-axis is the average individual-level PGS accuracy across 100 simulation replicates, and the error bars represent ±1.96 s.e.m. **d**, Population-level metrics of PGS accuracy recapitulates the decay in PGS accuracy across the genetic continuum. All UKBB testing individuals are divided into 100 equal-interval bins based on their GD. The *x*-axis is the average GD for the bin, and the *y*-axis is the squared correlation between genetic liability and PGS estimates for the individuals within the bin. The dot and error bars represent the mean and ±1.96 s.e.m from 100 simulations, respectively.

Next, we investigated the impact of GD on individual-level PGS accuracy. As expected, the width of the credible interval increases linearly with GD, reflecting reduced predictive accuracy for the PGS (Fig. 2b). The average width of the 90% credible interval is 1.83 in the furthest decile of GD, a 1.8-fold increase over the average width in the closest decile of GD. In contrast to the credible interval width, the individual-level PGS accuracy \(\hat{{r}_{i}^{2}}\) decreases with GD from the training data (Fig. 2c); the average estimated accuracy of individuals in the closest decile GD is fourfold higher than that of individuals in the furthest decile. Even among the most homogeneous grouping of individuals traditionally labelled as WB, we observe a 5% relative decrease in accuracy for individuals at the furthest decile of GD as compared to those in the closest decile. Similar results are observed when using a population-level PGS metric of accuracy, albeit at the expense of binning individuals according to GD; we find a high degree of concordance between the average \(\hat{{r}_{i}^{2}}\) within the bin and the population-level R^2 estimated within the bin (Fig. 2d and Extended Data Fig. 1a). Similarly, we observe a high consistency between average \(\hat{{r}_{i}^{2}}\) and squared correlation between PGS and simulated phenotypes ($R = 0.86$, $P < 10^{-10}$; Extended Data Fig. 1b). Taken together, our results show that the 90% credible intervals remain calibrated for individuals that are genetically distant from the training population at the expense of

wider credible intervals, and $\hat{{r}_{i}^{2}}$ captures the PGS accuracy decay across GD.

To demonstrate that the continuous accuracy decay is not specific to PGS models trained on European ancestries, we conducted further analyses using a non-European training dataset composed of individuals of NG and CB GIAs (we grouped the two GIAs to attain sufficient sample size for simulations). We simulated a high signal-to-noise trait by setting ${h}_{g}^{2}=0.8$ and proportion of causal variants ${p}_{{\rm{causal}}}=1\%$ and 0.1% with 56,539 SNPs on chromosome 10 alone. We trained PGS models on 5,000 individuals from the NG and CB GIA clusters and applied the models to the remaining testing individuals. The coverage of the 90% credible intervals was invariant to GD despite slight miscalibration. The 90% credible interval width increased and individual PGS accuracy decreased when the testing individual was further away from the training data. This trend is consistent with the observed decrease in empirical accuracy computed as squared correlation between PGS and genetic value as GD increases (Extended Data Figs. 2 and 3).

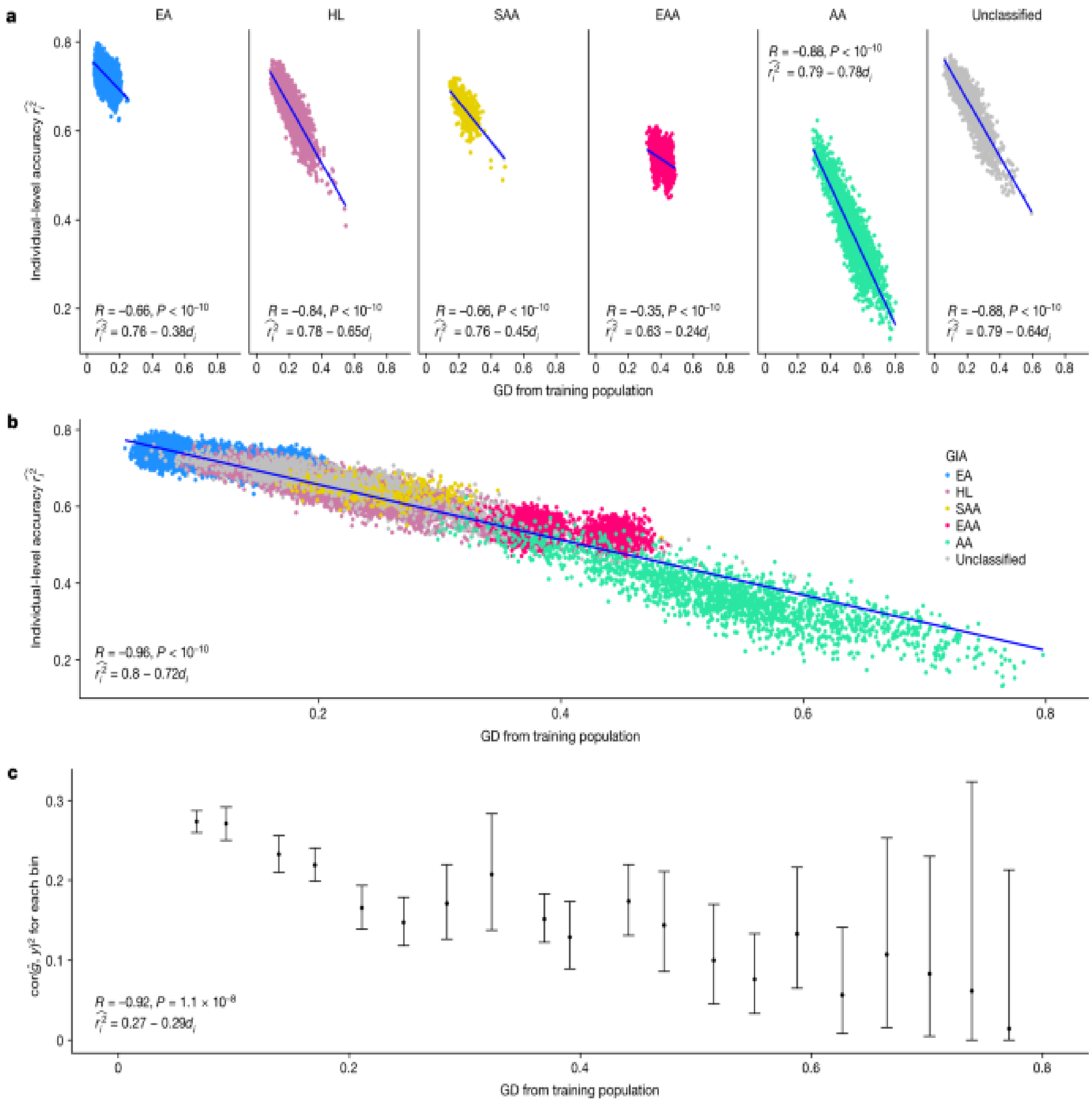
We further evaluated the impact of the number of PCs used for calculating GD on its ability to capture accuracy decay. We varied the number of PCs (J) from 1 to 20 and observed that the correlation between GD and individual accuracy ($-{\rm{cor}}({d}_{i}, {r}_{i}^{2}({g}_{i}, {\hat{g}}_{i}))$) increases when more PCs are used for computing GD, but no further improvement is observed when $J > 15$ for any GIA clusters or the whole biobank (Extended Data Fig. 4). Therefore, we set $J = 20$ for simplicity. We also explored average squared genetic relationship from training data as an alternative metric of GD and found that it is a better prediction of accuracy decay within each GIA clusters (Extended Data Fig. 4). However, because this metric relies on individual-level training data that are usually not available, we choose to use PCA-based GD for convenience.

PGS accuracy varies across the genetic continuum

Having validated our approach in simulations, we next turn to empirical data. For illustration purposes, we use height as an example, focusing on the ATLAS biobank as the target population with PGS trained on the 371,018

WB individuals from the UKBB (Methods); other traits show similar trends and are presented in the next sections. PGS accuracy at the individual level varies with GD across the entire biobank as well as within each GIA cluster (Fig. 3 and Extended Data Fig. 5). For example, GD strongly correlates with PGS accuracy of individuals in the GIA cluster labelled as Hispanic Latino American (HL, $R = -0.84$) and African American (AA, $R = -0.88$) in ATLAS. Notably, GD correlates with PGS accuracy even in non-admixed GIA clusters with correlations as −0.66, −0.66 and −0.35, for European American (EA), South Asian American (SAA) or East Asian American (EAA) GIA clusters, respectively. Similar qualitative results are also observed when applying PGS to test data from the UKBB; significant negative correlations are found between GD and individual PGS accuracy in all of the subcontinental GIA clusters in the UKBB (Extended Data Fig. 5), with correlation coefficients ranging from $R = -0.031$ for the WB cluster to $R = -0.62$ for the CB cluster.

Fig. 3: The individual-level accuracy for height PGS decreases across the genetic ancestry continuum in ATLAS.



a, Individual PGS accuracy decreases within both homogeneous and admixed genetic GIA clusters. Each dot represents a testing individual from ATLAS. For each dot, the *x*-axis represents its distance from the training population on the genetic continuum; the *y*-axis represents its PGS accuracy. The colour represents the GIA cluster. **b**, Individual PGS accuracy decreases across the entire ATLAS. **c**, Population-level PGS accuracy decreases with the average GD in each GD bin. All ATLAS individuals are divided into 20 equal-interval GD bins. The *x* axis is the average GD within the bin, and the *y* axis is the squared correlation between PGS and phenotype for individuals in the bin; the dot and error bar show the mean and 95% confidence interval from 1,000 bootstrap samples. *R* and *P* refer to the correlation between GD

and PGS accuracy and its significance, respectively, from two-sided Pearson correlation tests without adjustment for multiple hypothesis testing. Any *P* value below 10^{-10} is shown as $P < {10}^{-10}$. EA, European American; HL, Hispanic Latino American; SAA, South Asian American; EAA, East Asian American; AA, African American.

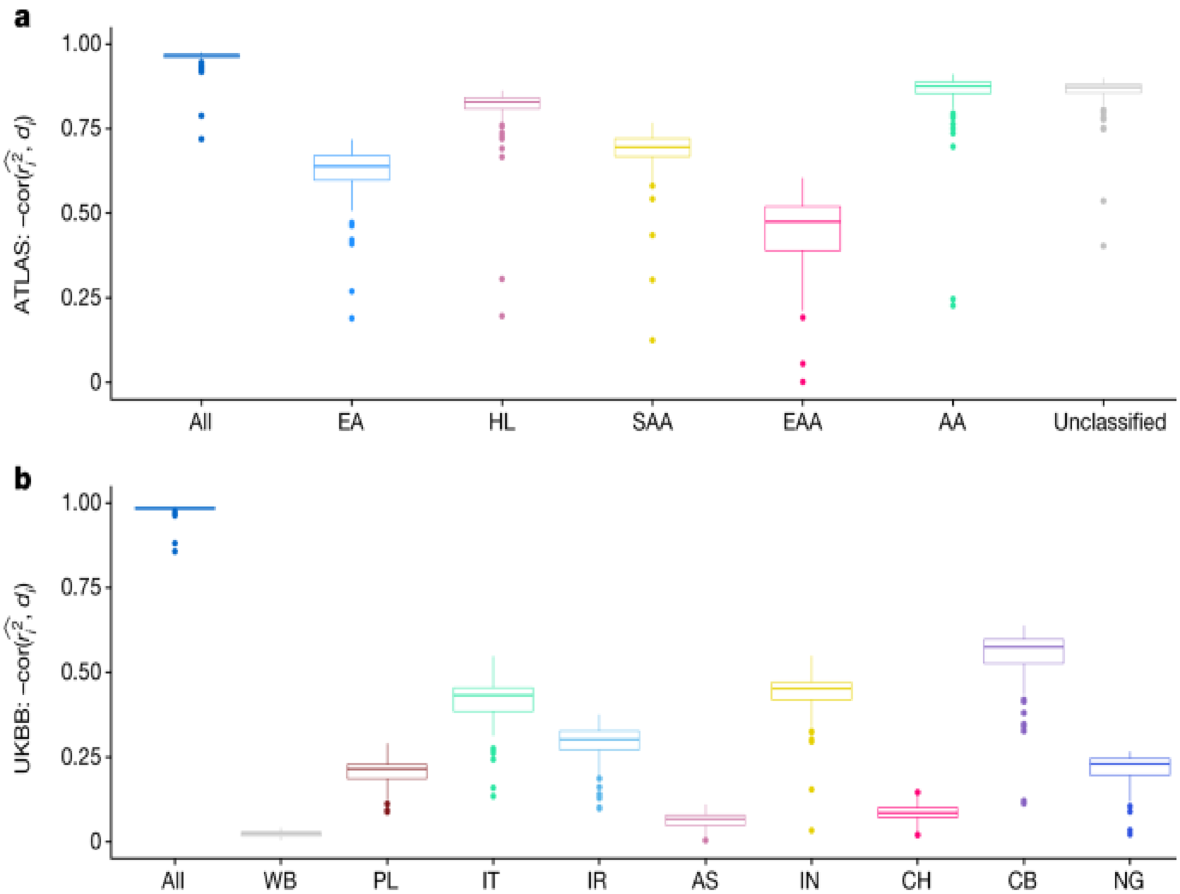
Next, we focused on the impact of GD on PGS accuracy across all ATLAS individuals regardless of GIA clustering ($R = -0.96$, $P < 10^{-10}$; Fig. 3b). Notably, we find a strong overlap of PGS accuracies across individuals from different GIA clusters demonstrating the limitation of using a single cluster-specific metric of accuracy. For example, when rank-ordering by GD, we find that the individuals from the closest GD decile in the HL cluster have similar estimated accuracy to the individuals from the furthest GD decile in EA cluster (average $\hat{{r}_{i}^{2}}$ of 0.71 versus 0.71). This shows that GD enables identification of HL individuals with similar PGS performance to the EA cluster thus partly alleviating inequities due to limited access to accurate PGS. Most notably, GD can be used to evaluate PGS performance for individuals that cannot be easily clustered by current genetic inference methods (6% of ATLAS; Fig. 3b) partly owing to limitations of reference panels and algorithms for assigning ancestries. Among this traditionally overlooked group of individuals, we find the GD ranging from 0.02 to 0.64 and their corresponding estimated PGS accuracy $\hat{{r}_{i}^{2}}$ ranging from 0.63 to 0.21. In addition to evaluating PGS accuracy with respect to the genetic liability, we also evaluated accuracy with respect to the residual height after regressing out sex, age and PC1–10 on the ATLAS from the actual measured trait. Using equally spaced bins across the GD continuum, we find that correlation between PGS and the measured height tracks significantly with GD ($R = -0.92$, $P = 1.1 \times 10^{-8}$; Fig. 3c).

PGS accuracy decay is pervasive

Having established the coupling of GD with PGS accuracy in simulations and for height, we next investigate whether this relationship is common across complex traits using PGSs for a broad set of 84 traits (Supplementary Table 1). We find consistent and pervasive correlations of GD with PGS accuracy across all considered traits in both ATLAS and the UKBB (Fig. 4).

For example, the correlations between GD and individual PGS accuracy range from −0.71 to −0.97 with an average of −0.95 across the 84 PGSs in ATLAS with similar results observed in the UKBB. Traits with sparser genetic architectures and fewer non-zero weights in the PGS have a lower correlation between GD and PGS accuracy; we reason that this is because GD represents genome-wide genetic variation patterns that may not reflect a limited number of causal SNPs well. For example, PGS for lipoprotein A (log_lipoA) has the lowest estimated polygenicity (0.02%) among the 84 traits and has the lowest correlation in ATLAS (−0.71) and the UKBB (−0.85). By contrast, we observe a high correlation between GD and PGS accuracy (>0.9) for all traits with an estimated polygenicity >0.1%. Next, we show that the fine-scale population structure accountable for the individual PGS accuracy variation is also prevalent within the traditionally defined genetic ancestry group. For example, in ATLAS we find that 501 of 504 (84 traits across 6 GIA clusters) trait–ancestry pairs have a significant associations between GD and individual PGS accuracy after Bonferroni correction. In the UKBB, we find 572 of the 756 (84 traits across 9 subcontinental GIA clusters) trait–ancestry pairs have significant associations between GD and PGS accuracy after Bonferroni correction. We also find that a more stringent definition of homogeneous GIA clusters results in a lower correlation magnitude (Extended Data Fig. 6). Empirical analyses of PGS accuracy show a similar trend. When averaging across 84 traits, we find that the empirical accuracy decreases with increased GD across GIA clusters as reported by previous studies[39]. Further analyses based on GD bins show the decreasing trend at a finer scale (Extended Data Fig. 7).

Fig. 4: The correlation between individual PGS accuracy and GD is pervasive across 84 traits across ATLAS and the UKBB.



a, The distribution of correlation between individual PGS accuracy and GD for 84 traits in ATLAS. **b**, The distribution of correlation between individual PGS accuracy and GD for 84 traits in the UKBB. Each box plot contains 84 points corresponding to the correlation between PGS accuracy and GD within the GIA group specified by the *x*-axis for each of the 84 traits. The box shows the first, second and third quartiles of the 84 correlations, and the whiskers extend to the minimum and maximum estimates located within 1.5 × IQR from the first and third quartiles, respectively. Numerical results are reported in Supplementary Tables 2 and 3.

PGS varies across the genetic continuum

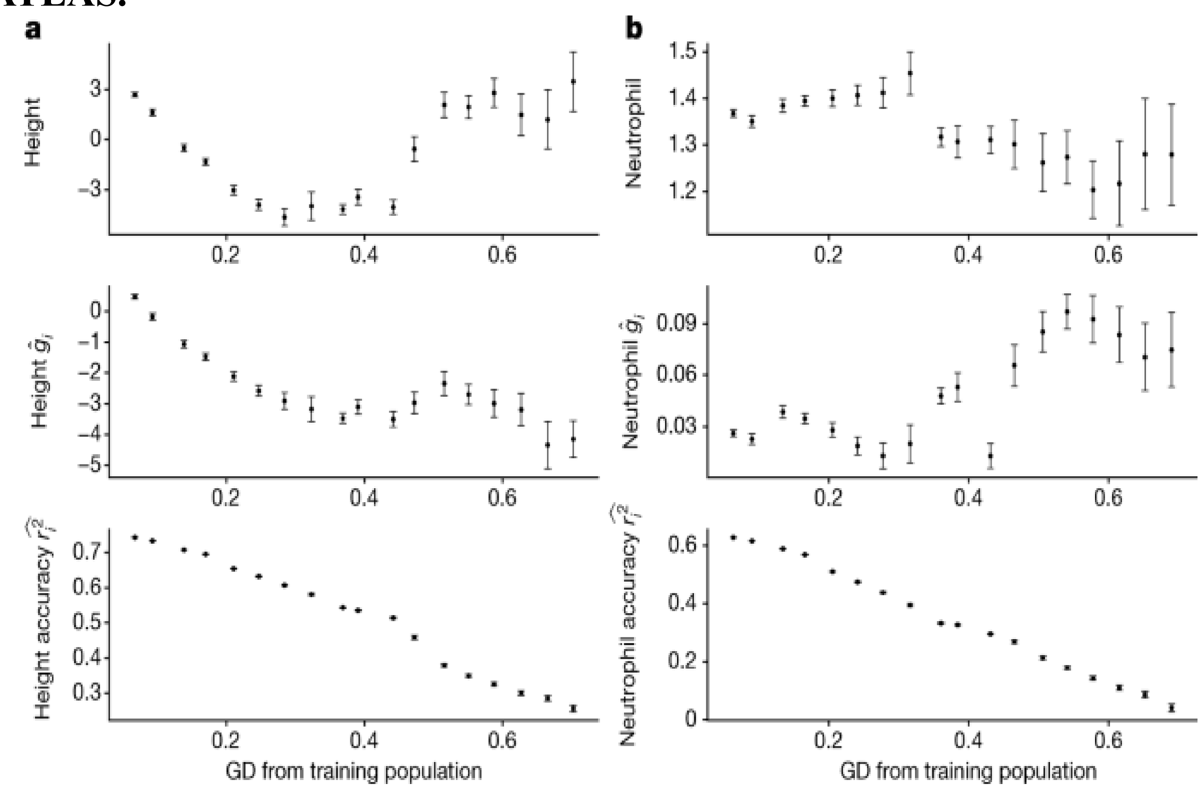
We have focused so far on investigating the relationship between GD (d_i) and PGS accuracy ($\hat{r}_i^2$). Next, we evaluate the impact of GD on PGS estimates ($\hat{g}_i$) themselves. We find a significant correlation between GD and PGS estimates for 82 of 84 traits,

with correlation coefficients ranging from $R = -0.52$ to $R = 0.74$ (Extended Data Fig. 8); this broad range of correlations is in stark contrast with the consistently observed negative correlation between GD and PGS accuracy. To better understand whether the coupling of PGS with GD is due to stratification or true signal, we compared the correlation of GD with PGS estimates ($\mathrm{cor}(d_i, \hat{g}_i)$) to the correlation of GD with measured phenotype values ($\mathrm{cor}(d_i, y_i)$). We find a wide range of couplings reflecting trait-specific signals; for 30 traits, GD correlates in opposite directions with PGS versus phenotype; for 40 traits, GD correlates in the same directions with PGS versus phenotype but differs in correlation magnitudes (Extended Data Fig. 8). For example, GD shows opposite and significantly different correlations for PGS versus trait for years of education (years_of_edu, $\mathrm{cor}(\,y_i, d_i) = 0.03$, $\mathrm{cor}(\,\hat{g}_i, d_i) = -0.18$). Other traits, such as hair colour, show a highly consistent impact of GD on PGS versus trait (darker_hair, $\mathrm{cor}(\,y_i, d_i) = 0.59$, $\mathrm{cor}(\,\hat{g}_i, d_i) = 0.74$), whereas for monocyte percentage, GD shows different magnitudes albeit with the same directions (monocyte_perc, $\mathrm{cor}(\,y_i, d_i) = -0.03$, $\mathrm{cor}(\,\hat{g}_i, d_i) = -0.52$). Moreover, GD correlates with PGS and phenotype even within the same GIA cluster and the correlation patterns vary across clusters (Extended Data Fig. 9).

The correlation of GD with phenotype and PGS is also observed in ATLAS. For example, both height phenotype and height PGS vary along GD in ATLAS (Fig. 5); this holds true even when restricting analysis to the EA genetic ancestry cluster (Supplementary Fig. 1). This is consistent with genetic liability driving difference in phenotypes but could also be explained by residual population stratification. For neutrophil counts, phenotype and PGS vary in opposite directions with respect to GD across the ATLAS (Fig. 5), although the trend is similar for phenotype and PGS in the EA GIA clusters (Supplementary Fig. 1). This could be explained by genetic liability driving signal in Europeans with stratification for other groups. Neutrophil counts have been reported to vary greatly across ancestry groups with reduced counts in individuals of African ancestries[54]. In ATLAS, we observe a negative correlation (−0.04) between GD and neutrophil counts in agreement with the previous reports, whereas GD is positively correlated

(0.08) with PGS estimates—genetically distant individuals traditionally labelled as African American having higher PGS than average. The opposite directions in phenotype–distance and PGS–distance correlations are partly attributed to the Duffy-null SNP rs2814778 on chromosome 1q23.2. This variant is strongly associated with neutrophil counts among individuals traditionally identified as African ancestry, but it is rare and excluded in our training data. This exemplifies the potential bias in PGS due to non-shared causal variants and emphasizes ancestral diversity in genetic studies. As PGS can vary across GD either as a reflection of true signal (that is, genetic liability varying with ancestry) or owing to biases in PGS estimation ranging from unaccounted residual population stratification to incomplete data (for example, partial ancestry-specific tagging of causal effects), our results emphasize the need to consider GD in PGS interpretation beyond adjusting for PGS \({r}_{i}^{2}\).

Fig. 5: Measured phenotype, PGS estimates and accuracy vary across ATLAS.



a, Variation of height phenotype, PGS estimates and accuracy across different GD bins in ATLAS. **b**, Variation of log neutrophil count phenotype, PGS estimates and accuracy across different GD bins in ATLAS. The 36,778

ATLAS individuals are divided into 20 equal-interval GD bins. Bins with fewer than 50 individuals are not shown owing to large s.e.m. All panels share the same layout: the *x* axis is the average GD within the bin; the *y* axis is the average phenotype (top), PGS (middle) and individual PGS accuracy (bottom); the error bars represent ±1.96 s.e.m.

Discussion

In this work, we have shown that PGS accuracy varies from individual to individual and proposed an approach to personalize PGS metrics of performance. We used a PCA-based GD[39] from the centre of training data to describe an individual's unique location on the genetic ancestry continuum and showed that individual PGS accuracy tracks well with GD. The continuous decay of PGS performance as the target individual becomes further away from the training population is pervasive across traits and ancestries. We highlight the variability in PGS performance along the continuum of genetic ancestries, even within traditionally defined homogeneous populations. As the genetic ancestries are increasingly recognized as continuous rather than discrete[7,42,43,44,45], the individual-level PGS accuracy provides a powerful tool to study PGS performance across diverse individuals to enhance the utility of PGS. For example, by using individual-level PGS accuracy, we can identify individuals from Hispanic Latino GIA who have similar PGS accuracy to individuals of European GIA, thus partly alleviating inequities due to lack of access to accurate PGS.

Simulation and real-data analyses show that individual PGS accuracy is highly correlated with GD, in alignment with existing works showing that decreased similarity (measured by relatedness, linkage disequilibrium and/or minor allele frequency differences, fixation index (F_{st}) and so on)[41,55] between testing individuals and training data is a major contributor to PGS accuracy decay. However, practical factors that may affect transferability, such as genotype–environment interaction and population-specific causal variants, are not modelled in the calculation of individual PGS accuracy and this is left for future work.

Our results emphasize the importance of PGS training in diverse ancestries[56] as it can provide advantages for all individuals. Broadening PGS training beyond European ancestries can lead to improved accuracy in genetic effect estimation particularly for variants with higher frequencies in non-European data. It can also increase PGS portability by reducing the GD from target to training data. However, increased diversity may also bring challenges to statistical modelling; for example, differences in genetic effects may correlate with environment factors and could bias genetic risk prediction. To address these challenges, more sophisticated statistical methods are needed that can effectively leverage ancestrally diverse populations to train PGS[3] (for example, PRS-CSx[57], vilma[58] and CT-SLEB[59]). Concerted global effort and equitable collaborations are also crucial to increase the sample size of underrepresented individuals as part of an effort to reduce health disparities across ancestries[56,60].

We highlight the pervasive correlation between PGS estimates and GD of varying magnitude and sign as compared to the correlation between phenotype and GD. This provides a finer resolution of the mean shift of PGS estimates across genetic ancestry groupings[38]. The correlation between GD and PGS estimates can arise from bias and/or true biological difference, and more effort is needed to investigate the PGS bias in the context of genetic ancestry continuum.

We note several limitations and future directions of our work. First, our proposed individual PGS accuracy is an upper bound of true accuracy and should be interpreted only in terms of the additive heritability captured by SNPs included in the model. Missing heritability[61,62] and misspecification of the heritability model along with population-specific causal variants and effect sizes may further decrease real accuracy. For example, the prediction accuracy for neutrophil count is overestimated among African American individuals because the Duffy-null SNP rs2814778 (ref. [54]) is not captured in the UKBB WB training data. Future work could investigate the impact of the population-specific components of genetic architecture on the calibration of PGS accuracy. Second, we approximate the variance of genetic liability in the denominator of equation (1) with heritability and set a fixed value for all individuals. Preliminary results show that replacing the denominator with a Monte Carlo estimation of genetic liability variance recapitulates the

accuracy decay in estimated PGS accuracy, albeit the correlation is slightly reduced (Extended Data Fig. 10). Third, individual PGS accuracy evaluates how well the PGS estimates the genetic liability instead of phenotype. Quantifying the individual accuracy of PGS with respect to phenotype can be achieved by also modelling non-genetic factors for proper calibration. Fourth, limited by sample size, we combined GIA groups as a training set in simulation experiments to replicate PGS accuracy decay; this is not an optimal strategy for data analysis as the population structure in the training data may confound the true genetic effects and reduce prediction accuracy. We leave a more comprehensive investigation of non-European PGS training data for future work. Sixth, although we advocate for the use of continuous genetic ancestry, we trained our PGS models on a discrete GIA cluster of WB because current PGS methods rely on discrete genetic ancestry groupings. We leave the development of PGS training methods that are capable of modelling continuous ancestries as future work. Finally, we highlight that, just like PGS, the traditional clinical risk assessment may suffer from limited portability across diverse populations[18]. For examples, the pooled cohort equation overestimates atherosclerotic cardiovascular disease risk among non-European populations[16]; and a traditional clinical breast cancer risk model developed in the European population in the USA overestimated the breast cancer risk among older Korean women[17]. Here we focus on genetic prediction potability owing to the wide interest and attention from both the research community and society. We emphasize that improving the portability of traditional clinical risk factor models in diverse populations is an essential component of health equity and requires thorough investigation.

Methods

Model

We model the phenotype of an individual with a standard linear model ${y}_{i}={x}_{i}^{{\rm{\top }}}\beta +{{\epsilon }}_{i}$, in which ${x}_{i}$ is an $M\times 1$ vector of standardized genotypes (centred and standardized with respect to the allele frequency in the training population for both training and testing individuals), β is an $M\times 1$ vector of standardized genetic effects, and ${{\epsilon }}_{i}$ is random noise.

Under a random effects model, β is a vector of random variable sampled from a prior distribution $p(\beta)$ that differs under different genetic architecture assumptions[62] and PGS methods[53,63,64,65]. The PGS weights $\hat{\beta}=E_{\beta|D}(\beta)$ are estimated to be the posterior mean given the observed data D ($D=\left(X_{\mathrm{train}},y_{\mathrm{train}}\right)$ with access to individual-level genotype, X_{train}, and phenotype, y_{train}; or $D=\left(\hat{\beta}_{\mathrm{GWAS}},\,\hat{R}\right)$ with access to marginal association statistics $\hat{\beta}_{\mathrm{GWAS}}$ and LD matrix $\hat{R}$, in which GWAS stands for genome-wide association study). The genetic liability ($g_i=x_i^T\beta$) of an individual *i* is estimated to be $\hat{g}_i=E_{\beta|D}(x_i^T\beta)$, the uncertainty of which is estimated as the posterior variance of genetic liability $\mathrm{var}(\,\hat{g}_i)=\mathrm{var}_{\beta|D}(x_i^T\beta)$ (ref. [49]).

Individual PGS accuracy

We define individual PGS accuracy as the squared correlation between an individual's genetic liability, g_i, and its PGS estimate, $\hat{g}_i$, following the general form in ref. [28]:

$$r_i^2=\frac{\mathrm{cov}_{\beta,D}(g_i,\hat{g_i})^2}{\mathrm{var}_{\beta,D}(\,g_i)\,\mathrm{var}_{\beta,D}(\,\hat{g_i})}=\frac{\mathrm{var}_D(x_i^{\top}\hat{\beta})^2}{\mathrm{var}_{\beta}(x_i^{\top}\beta)\,\mathrm{var}_D(x_i^{\top}\hat{\beta})}$$

Here we are interested in the PGS accuracy of a given individual; therefore, the genotype is treated as a fixed variable, and genetic effects are treated as a random variable. We note that a random effects model is essential; otherwise, $\mathrm{cov}_{\beta,D}\left(g_i\,,\hat{g_i}\right)$ and $\mathrm{var}_{\beta,D}\left(\,g_i\right)$ are 0. Under a random effects model, both the genetic liability and PGS estimate for individual i

are random variables. The randomness of ${g}_{i}={x}_{i}^{{\rm{\top }}}\beta$ comes from the randomness in β, and the randomness of ${\hat{g}}_{i}={x}_{i}\hat{\beta }$ comes from the randomness of both β and the training data D. Individual PGS accuracy measures the correlation between ${g}_{i}$ and ${\hat{g}}_{i}$, which can be computed with the following equation:

$${r}_{i}^{2}=1-\frac{{E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)}$$
(2)

in which ${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta)$ is the posterior variance of genetic liability given the training data, and ${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ is the genetic variance. The equation is derived as follows.

First, we show that under the random effects model, ${{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}({x}_{i}^{{\rm{\top }}}\hat{\beta },\,{x}_{i}^{{\rm{\top }}}\beta)={{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({x}_{i}^{{\rm{\top }}}\hat{\beta })$ (in which $\hat{\beta }={E}_{\beta {|D}}(\beta)$) following equation 5.149 in ref. [66]:

$$\begin{array}{c}{{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}(\hat{\beta },{\beta }^{{\rm{\top }}})=\,{E}_{\beta ,D}(\hat{\beta }{\beta }^{{\rm{\top }}})-{E}_{\beta ,D}(\hat{\beta }){E}_{\beta ,D}({\beta }^{{\rm{\top }}})\\ \,\,=\,{E}_{D}({E}_{\beta |D}(\hat{\beta }{\beta }^{{\rm{\top }}}))-{E}_{D,\beta }(\hat{\beta }){E}_{D}({E}_{\beta |D}({\beta }^{{\rm{\top }}}))\\ \,\,=\,{E}_{D}({E}_{\beta |D}({E}_{\beta |D}(\beta){\beta }^{{\rm{\top }}}))-\,{E}_{D}({E}_{\beta |D}(\beta)){E}_{D}({E}_{\beta |D}({\beta }^{{\rm{\top }}}))\\ \,\,=\,{E}_{D}({E}_{\beta |D}(\beta){E}_{\beta |D}({\beta }^{{\rm{\top }}}))-\,{E}_{D}({E}_{\beta |D}(\beta)){E}_{D}({E}_{\beta |D}({\beta }^{{\rm{\top }}}))\\ \,\,=\,{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({E}_{\beta |D}(\beta))\\ \,\,=\,{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}(\hat{\beta })\end{array}$$

Multiplying ${x}_{i}$ on both sides of the equation, we obtain:

$${x}_{i}^{T}{{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}(\hat{\beta },\,\beta){x}_{i}={x}_{i}^{T}{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}(\hat{\beta }){x}_{i}$$

$${{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}({x}_{i}^{{\rm{\top }}}\hat{\beta },\,{x}_{i}^{{\rm{\top }}}\beta)={{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({x}_{i}^{{\rm{\top }}}\hat{\beta })$$

(3)

Equation (3) also implies the slope from regression of observed phenotypic values (or true genetic liability) on the estimated PGS equal to 1 (Supplementary Fig. 2), which offers an alternative way to assess the calibration of PGS as done in refs. [64,67].

$${\rm{s}}{\rm{l}}{\rm{o}}{\rm{p}}{\rm{e}}=\frac{{\rm{c}}{\rm{o}}{\rm{v}}({x}_{i}^{{\rm{\top }}}\hat{\beta }\,,\,{y}_{i})}{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\hat{\beta })}=\frac{{\rm{c}}{\rm{o}}{\rm{v}}({x}_{i}^{{\rm{\top }}}\hat{\beta }\,,\,{x}_{i}^{{\rm{\top }}}\beta +{{\epsilon }}_{i})}{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\hat{\beta })}=\frac{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\hat{\beta })}{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\hat{\beta })}=1$$

Next, by applying the law of total variance, we show that:

$${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}({g}_{i})={{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}({x}_{i}^{T}\beta)={E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{T}\beta))+{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({E}_{\beta |D}({x}_{i}^{T}\beta))$$

$${{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({x}_{i}^{{\rm{\top }}}\hat{\beta })={{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}({x}_{i}^{{\rm{\top }}}\beta)-\,{E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))$$

(4)

Third, we derive the correlation between ${g}_{i}$ and ${\hat{g}}_{i}$ as:

$$\begin{array}{c}{r}_{i}^{2}=\frac{{{\rm{c}}{\rm{o}}{\rm{v}}}_{\beta ,D}{({g}_{i},{\hat{g}}_{i})}^{2}}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}(\,{g}_{i}){{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}(\,{\hat{g}}_{i})}\\ \,=\,\frac{{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}{({x}_{i}^{{\rm{\top }}}\hat{\beta })}^{2}}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta){{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({x}_{i}^{{\rm{\top }}}\hat{\beta })}\,{\rm{b}}{\rm{y}}\,{\rm{a}}{\rm{p}}{\rm{p}}{\rm{l}}{\rm{y}}{\rm{i}}{\rm{n}}{\rm{g}}\,{\rm{e}}{\rm{q}}{\rm{u}}{\rm{a}}{\rm{t}}{\rm{i}}{\rm{o}}{\rm{n}}\,(3)\\ \,=\,\frac{{{\rm{v}}{\rm{a}}{\rm{r}}}_{D}({x}_{i}^{{\rm{\top }}}\hat{\beta })}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)}\\ \,=\,\frac{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta ,D}({x}_{i}^{{\rm{\top }}}\beta)-\,{E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)}\,{\rm{b}}{\rm{y}}\,{\rm{a}}{\rm{p}}{\rm{p}}{\rm{l}}{\rm{y}}{\rm{i}}{\rm{n}}{\rm{g}}\,{\rm{e}}{\rm{q}}{\rm{u}}{\rm{a}}{\rm{t}}{\rm{i}}{\rm{o}}{\rm{n}}\,(4)\\ \,=\,1-\frac{\,{E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))}{{{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)}\end{array}$$

The above equation is widely used in animal breeding theory to compute the reliability of estimated breeding value for each individual[30]. In this work, we use individual PGS uncertainty $({\rm{v}}{\rm{a}}{\rm{r}}(\,{\hat{g}}_{i})={{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta)$ as an unbiased estimator of ${E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))$. We also use estimated heritability to approximate ${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ in simulations in which the phenotype has unit variance. In real-data analysis, as the phenotype does not necessarily have unit variance, we approximate ${{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ by scaling the estimated heritability with the residual phenotypic variance in the training population after regressing GWAS covariates including sex, age and precomputed UKBB PC1–16 (Data-Field 22009).

Analytical form of individual PGS accuracy under infinitesimal assumption

Without loss of generality, we assume a prior distribution of genetic effects as follows:

$$p\left(\beta |{\sigma }_{\beta }^{2}\right)={MVN}(0,{\sigma }_{\beta }^{2}{I}_{M})$$

where *M* is the number of genetic variants. With access to individual genotype, \({X}_{{\rm{train}}}\), and phenotype, \({y}_{{\rm{train}}}\), data, the likelihood of the data is

$$p(\,{y}_{{\rm{train}}}\,{\rm{| }}\,{X}_{{\rm{train}}},\beta \,,\,{\sigma }_{{\rm{e}}}^{2})=MVN({X}_{{\rm{train}}}\,\beta \,,\,{\sigma }_{{\rm{e}}}^{2}{I}_{N})$$

where *N* is the training sample size. The posterior distribution of genetic effects given the data is proportional to the product of the prior and the likelihood:

$$\begin{array}{cc}p(\beta |{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{{\sigma }}_{\beta }^{2},{{\sigma }}_{{\rm{e}}}^{2}) & \propto \,p(\beta |{{\sigma }}_{\beta }^{2})p(\,{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}|{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},\beta ,{{\sigma }}_{{\rm{e}}}^{2})\\ & \propto \,MVN(0,{{\sigma }}_{\beta }^{2}{I}_{M})MVN({X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\,\beta ,{{\sigma }}_{{\rm{e}}}^{2}{I}_{N})\\ & \propto \,MVN({\mu }_{\beta },{{\sigma }}_{\beta })\end{array}$$

in which \({\mu }_{\beta }={\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}{I}_{M}+{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\right)}^{-1}{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}\,{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\) and \({\varSigma }_{\beta }={{\sigma }}_{{\rm{e}}}^{2}\,{\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}

{I}_{M}+{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\right)}^{-1}\). This form is equivalent to the solution of random effects in the best linear unbiased prediction with the pedigree matrix or genetic relationship matrix[29,32].

For a new target individual, the posterior variance of the genetic liability is:

$${\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\beta \,|\,{x}_{i},{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{{\sigma }}_{\beta }^{2},{{\sigma }}_{{\rm{e}}}^{2})={x}_{i}^{{\rm{\top }}}{\varSigma }_{\beta }{x}_{i}={{\sigma }}_{{\rm{e}}}^{2}\,{x}_{i}^{{\rm{\top }}}\,{\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}{I}_{M}+{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\right)}^{-1}{x}_{i}$$

After carrying out eigendecomposition on ${X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}\,{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}={\sum }_{j=1}^{J}{\lambda }_{j}{v}_{j}{v}_{j}^{{\rm{\top }}}$, we can rewrite

$${\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}{I}_{M}+{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\right)}^{-1}={\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}{I}_{M}+\mathop{\sum }\limits_{j=1}^{J}{\lambda }_{j}{v}_{j}{v}_{j}^{{\rm{\top }}}\right)}^{-1}=\mathop{\sum }\limits_{j=1}^{J}{\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}+{\lambda }_{j}\right)}^{-1}{v}_{j}{v}_{j}^{{\rm{\top }}}$$

in which ${\lambda }_{j}$ and ${v}_{j}$ correspond to the *j*th eigenvalue and unit-length eigenvector of the training genotype, ${X}_{{\rm{train}}}$.

Thus, we can rewrite the posterior variance of genetic liability as

$${\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\beta |{x}_{i},{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{{\sigma }}_{\beta }^{2},{{\sigma }}_{{\rm{e}}}^{2})=\,{{\sigma }}_{{\rm{e}}}^{2}\mathop{\sum }\limits_{j=1}^{J}{\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}+{\lambda }_{j}\right)}^{-1}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}$$

Replacing \({E}_{D}({{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta))\) in equation (2) with the analytical form of \({\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\beta |{x}_{i},{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{{\sigma }}_{\beta }^{2},{{\sigma }}_{e}^{2})\), we get

$${r}_{i}^{2}=1-\frac{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\beta |{x}_{i},{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{y}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}},{{\sigma }}_{\beta }^{2},{{\sigma }}_{{\rm{e}}}^{2})}{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{i}^{{\rm{\top }}}\beta)}=1-\frac{{{\sigma }}_{{\rm{e}}}^{2}{\sum }_{j=1}^{J}{\left(\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}+{\lambda }_{j}\right)}^{-1}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}}{{{\sigma }}_{\beta }^{2}{x}_{i}^{{\rm{\top }}}{x}_{i}}$$

As the eigenvalue of \({X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}^{{\rm{\top }}}\,{X}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}\) increases linearly with training sample size *N* (ref. [68]), at the UKBB-level sample size (for example, *N* = 371,018 for our UKBB WB training data), the eigenvalues for the top PCs are usually larger than the ratio of environmental noise variance and genetic variance \(\frac{{\sigma }_{e}^{2}}{{{\sigma }}_{\beta }^{2}}\). Thus, we can further approximate the analytical form with:

$${r}_{i}^{2}=1-\frac{{{\sigma }}_{{\rm{e}}}^{2}{\sum }_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{T}v}_{j}{v}_{j}^{T}{x}_{i}}{{{\sigma }}_{\beta }^{2}{x}_{i}^{T}{x}_{i}}=1-\frac{{{\sigma }}_{{\rm{e}}}^{2}}{{{\sigma }}_{\beta }^{2}}\frac{{\sum

}_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{T}v}_{j}{v}_{j}^{T}{x}_{i}}{{x}_{i}^{T}{x}_{i}}$$

The term ${\sum }_{j=1}^{J}\frac{1}{{\lambda }_{j}}{{x}_{i}^{{\rm{\top }}}v}_{j}{v}_{j}^{{\rm{\top }}}{x}_{i}$ is the squared Mahalanobis distance of the testing individual i from the centre of the training genotype data on its PC space and ${x}_{i}^{{\rm{\top }}}{x}_{i}$ is the sum of squared genotype across all variants. Empirically, the ratio between the two is highly correlated with the Euclidean distance of the individual from the training data on that PC space ($R = 1$, P value $< 2.2 \times 10^{-16}$ in the UKBB).

Genetic distance (GD)

The GD is defined as the Euclidean distance between a target individual and the centre of training data on the PC space of training data.

$${d}_{i}=\sqrt{{\sum }_{j=1}^{J}{({x}_{i}^{{\rm{\top }}}{v}_{j}-{\bar{x}}_{{\rm{t}}{\rm{r}}{\rm{a}}{\rm{i}}{\rm{n}}}{v}_{j})}^{2}}=\sqrt{{\sum }_{j=1}^{J}{({x}_{i}^{{\rm{\top }}}{v}_{j})}^{2}}$$

in which ${d}_{{i}}$ is the GD of a testing individual i from the training data, ${x}_{i}$ is an $M\times 1$ standardized genotype vector for testing individual i, ${v}_{{j}}$ is the jth eigenvector for the genotype matrix of training individuals, ${\bar{x}}_{{\rm{train}}}$ is the average genotype in the training population (${\bar{x}}_{{\rm{train}}}{v}_{j}=0$ given that the genotypes are centred with respect to the allele frequency in the training population), and J is set to 20.

Ancestry ascertainment in UKBB

The UKBB individuals are clustered into nine subcontinental GIA clusters—WB (white British), PL (Poland), IR (Iran), IT (Italy), AS (Ashkenazi), IN (India), CH (China), CB (Caribbean) and NG (Nigeria)—based on the top 16 precomputed PCs (Data-Field 22009) as described in ref. [39]. First, UKBB participants are grouped by country of origin (Data-Field 20115) and the

centre of each country on the PC space is computed as the geometric median for all countries, which serves as a proxy for the centre for each subcontinental ancestry. The centre of Ashkenazi GIA is determined using a dataset from ref. [69]. Second, we reassign each individual to one of the nine GIA groups on the basis of their Euclidean distance to the centres on the PC space, as the self-reported country of origin does not necessarily match an individual's genetic ancestry. The genetic ancestry of an individual is labelled as unknown if its distance to any genetic ancestry centre is larger than one-eighth of the maximum distance between any pairs of subcontinental ancestry clusters. We are able to cluster 91% of the UKBB participants into 411,018 WB, 4,127 PL, 1,169 IR, 6,499 IT, 2,352 AS, 1,798 CH, 2,472 CB and 3,894 NG. GIAs are not necessarily reflective of the full genetic diversity of a particular region but reflect only the diversity present in the UKBB individuals.

Ancestry ascertainment in ATLAS

The ATLAS individuals are clustered into five GIA clusters—European Americans (EA), Hispanic Latino Americans (HL), South Asian Americans (SAA), East Asian Americans (ESA) and African Americans (AA)—as described in ref. [5] on the basis of their proximity to 1000 Genome super populations on the PC space. First, we filter the ATLAS-typed genotypes with plink2 by Mendel error rate (plink --me 1 1 –set-me-missing), founders (--filter-founders), minor allele frequency (–maf 0.15), genotype missing call rate (--geno 0.05) and Hardy–Weinberg equilibrium test *P* value (–hwe 0.001). Next, ATLAS genotypes were merged with the 1000 Genomes phase 3 dataset. Then, linkage disequilibrium (LD) pruning was carried out on the merged dataset (--indep 200 5 1.15 --indep-pairwise 100 5 0.1). The top 10 PCs were computed with the flashpca2 (ref. [70]) software with all default parameters. Next, we use the super population label and PCs of the 1000 Genome individuals to train the *K*-nearest neighbours model to assign genetic ancestry labels to each ATLAS individual. For each ancestry cluster, we run the *K*-nearest neighbours model on the pair of PCs that capture the most variation for each genetic ancestry group: the European, East Asian and African ancestry groups use PCs 1 and 2, the Admixed American group uses PCs 2 and 3, and the South Asian group uses PCs 4 and 5. In each analysis, we use tenfold cross-validation to select the *k* hyper-parameter

from $k = 5$, 10, 15, 20. If an individual is assigned to multiple ancestries with probability larger than 0.5 or is not assigned to any clusters, their ancestry is labelled as unknown. We label the five 1000 Genome super population as EA for Europeans, HL for Admixed Americans, SAA for South Asians, AA for Africans and ESA for East Asians. We can cluster 95% of the ATLAS participants into 22,380 EA, 6,973 HL, 625 SAA, 3,331 EAA and 1,995 AA, and the ancestry of 2,332 individuals is labelled as unknown.

Genotype data

In simulations, we use 1,054,151 UKBB HapMap 3 SNPs for simulating phenotypes, training PGS models and calculating PGS for testing individuals in UKBB. For real-data analysis, we use an intersection of UKBB HapMap 3 SNPs and ATLAS imputed SNPs for the training of PGS in UKBB and calculating PGS for remaining UKBB individuals and ATLAS individuals. We start from 1,054,151 UKBB HapMap 3 SNPs and 8,048,268 ATLAS imputed SNPs. As UKBB is on genome build hg37 and ATLAS is on hg38, we first lift all ATLAS SNPs from hg38 to hg37 with the snp_modifyBuild function in the bigsnpr R package. Next, we match UKBB SNPs and ATLAS SNPs by chromosome and position with the snp_match function in bigsnpr. Then, we recode ATLAS SNPs using UKBB reference alleles with the plink2 --recode flag. In the end, 979,457 SNPs remain for training the LDpred2 models in real-data analysis.

Simulated phenotypes

We use simulations on all UKBB individuals to investigate the impact of GD from training data on the various metrics of PGS. We fix the proportion of causal SNPs $p_{\rm causal}=0.01$ and heritability as $h_g^2=0.25$. The simulated genetic effects and phenotype are generated as follows. First, we randomly sample

$${\beta }_{m}\sim \{\begin{array}{cc}N\left(0,\frac{{h}_{g}^{2}}{{\rm{v}}{\rm{a}}{\rm{r}}({x}_{m})M{p}_{{\rm{c}}{\rm{a}}{\rm{u}}{\rm{s}}{\rm{a}}{\rm{l}}}}\right) & {c}_{j}=1,{\rm{w}}{\rm{i}}{\rm{t}}{\rm{h}}\,{\rm{p}}{\rm{r}}{\rm{o}}{\rm{b}}{\rm{a}}{\rm{b}}{\rm{i}}{\rm{l}}{\rm{i}}{\rm{t}}{\rm{y}}\,{p}_{{\rm{c}}{\rm{a}}}

{\rm{u}}{\rm{s}}{\rm{a}}{\rm{l}}}\\ 0 & \,{c}_{j}=0,{\rm{w}}{\rm{i}}{\rm{t}}{\rm{h}}\,{\rm{p}}{\rm{r}}{\rm{o}}{\rm{b}}{\rm{a}}{\rm{b}}{\rm{i}}{\rm{l}}{\rm{i}}{\rm{t}}{\rm{y}}\,1-{p}_{{\rm{c}}{\rm{a}}{\rm{u}}{\rm{s}}{\rm{a}}{\rm{l}}}\end{array}$$

in which $\mathrm{var}({x}_{m})$ is the variance of allele counts for SNP m among all UKBB individuals. Second, we compute the genetic liability for each individual as ${g}_{i}={\sum }_{m=1}^{M}{x}_{im}{\beta }_{m}$ and randomly sample environmental noise ${{\epsilon }}_{i}\sim N(0,1-{h}_{g}^{2})$. Third, we generate phenotype as ${y}_{i}={g}_{i}+{{\epsilon }}_{i}$. We repeat the process 100 times to generate 100 sets of genetic liability and phenotypes.

GD from PGS training data

To compute the GD of testing individuals from the training population, we carry out PCA on the 371,018 UKBB WB training individuals and project the 48,586 UKBB testing individuals and 36,778 ATLAS testing individuals on the PC space. We start from the 979,457 SNPs that are overlapped in UKBB and ATLAS. First, we carry out LD pruning with plink2 (--indep-pairwise 1000 50 0.05) and exclude the long-range LD regions. Next, we carry out PCA analysis with flashpca2 (ref. [70]) on the 371,018 UKBB WB training individuals to obtain the top 20 PCs. Then, we project the remaining 48,586 UKBB individuals that are not included in the training data and 36,778 ATLAS individuals onto the PC space of training data by using SNP loadings (--outload loadings.txt) and their means and standard deviations (--outmeansd meansd.txt) output from flashpca2. In the end, we compute the GD for each individual as the Euclidean distance of their PCs from the centre of training data with the equation ${d}_{i}=\sqrt{{\sum }_{j=1}^{20}{\left({{\rm{pc}}}_{ij}\right)}^{2}}$, in which ${{\rm{pc}}}_{ij}$ is the ${j}_{{th}}$ PC of individual i.

LDpred2 PGS model training

The PGS models were trained on 371,018 UKBB individuals labelled as WB with the LDpred2 (ref. [53]) method for both simulation and real-data analysis. For simulation analysis, we use 1,054,151 UKBB HapMap 3 variants. For

real-data analysis, we use 979,457 SNPs that are overlapped in UKBB HapMap 3 variants and ATLAS imputed genotypes.

First, we obtain GWAS summary statistics by carrying out GWAS on the training individuals with plink2 using sex, age and precomputed PC1–16 as covariates. Second, we calculate the in-sample LD matrix with the function snp_cor from the R package bigsnpr[71]. Next, we use the GWAS summary statistics and LD matrix as input for the snp_ldpred2_auto function in bigsnpr to sample from the posterior distribution of genetic effect sizes. Instead of using a held-out validation dataset to select hyperparameters p (proportion of causal variants) and $h2$ (heritability), snp_ldpred2_auto estimates the two parameters from data with the Markov chain Monte Carlo (MCMC) method directly. We run 10 chains with different initial sparsity p from 10^{-4} to 1 equally spaced in log space. For all chains, we set the initial heritability as the LD score regression heritability[72] estimated by the built-in function snp_ldsc. We carry out quality control of the 10 chains by filtering out chains with estimated heritability that is smaller than 0.7 times the median heritability of the 10 chains or with estimated sparsity that is smaller than 0.5 times the median sparsity or larger than 2 times the median sparsity. For each chain that passes filtering, we remove the first 100 MCMC iterations as burn-in and thin the next 500 iterations by selecting every fifth iteration to reduce autocorrelation between MCMC samples. In the end, we obtain an $M \times B$ matrix \([\,{\widetilde{\beta }}^{(1)},\,{\widetilde{\beta }}^{(2)},\,...,\,{\widetilde{\beta }}^{(B)}]\), in which each column of the matrix \({\widetilde{\beta }}^{(b)}\) is a sample of posterior causal effects of the M SNPs. Owing to the quality control of MCMC chains, the total number of posterior samples \(B\) ranges from 500 to 1,000.

Calculation of PGS and accuracy

We use the score function in plink2 to compute the PGS for 48,586 and 36,778 testing individuals in UKBB and ATLAS, respectively. For each \({\widetilde{\beta }}^{(b)}\), we compute the PGS for each individual i as \({x}_{i}^{{\rm{\top }}}{\mathop{\beta }\limits^{ \sim }}^{(b)}\) with plink2 (--score). For each individual with genotype \({x}_{i}\), we compute \({{x}_{i}^{{\rm{\top }}}\mathop{\beta }\limits^{ \sim }}^{(1)},\,{{x}_{i}^{{\rm{\top }}}\mathop{\beta }\limits^{ \sim }}^{(2)}\,,\,...,\,

${{x}_{i}^{{\rm{\top }}}\mathop{\beta }\limits^{ \sim }}^{(B)}$ to approximate its posterior distribution of genetic liability. The genotype ${x}_{i}^{{\rm{\top }}}$ is centred to the average allele count (--read-freq) in training data to reduce the uncertainty from the unmodelled intercept. We estimate the PGS with the posterior mean of the genetic liability as ${\hat{g}}_{i}\,=\,{E}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta)\,=\,\frac{1}{B}{\sum }_{b=1}^{B}{x}_{i}^{{\rm{\top }}}{\mathop{\beta }\limits^{ \sim }}^{(b)}$. We estimate the individual-level PGS uncertainty as ${\rm{v}}{\rm{a}}{\rm{r}}({\hat{g}}_{i})={{\rm{v}}{\rm{a}}{\rm{r}}}_{\beta |D}({x}_{i}^{{\rm{\top }}}\beta)\,=$$\frac{1}{B}{\sum }_{b=1}^{B}{({x}_{i}^{{\rm{\top }}}{\mathop{\beta }\limits^{ \sim }}^{(b)}-{\hat{g}}_{i})}^{2}$. The individual-level PGS accuracy is calculated as $\hat{{r}_{i}^{2}}=1-\frac{{\rm{v}}{\rm{a}}{\rm{r}}({\hat{g}}_{i})}{{h}_{g}^{2}}$ for simulation (${h}_{g}^{2}$ is the heritability estimated by the LDpred2 model) and as $\hat{{r}_{i}^{2}}=1-\frac{{\rm{v}}{\rm{a}}{\rm{r}}({\hat{g}}_{i})}{{h}_{g}^{2}{\rm{v}}{\rm{a}}{\rm{r}}({y}_{train}-{\hat{y}}_{train})}$ for real-data analysis, in which ${\rm{v}}{\rm{a}}{\rm{r}}(\,{y}_{{\rm{train}}}-{\hat{y}}_{{\rm{train}}})$ is the variance of residual phenotype in training data after regressing out GWAS covariates.

Calibration of credible interval in simulation

We run the LDpred2 model on 371,018 WB training individuals for the 100 simulation replicates. In each simulation, for individual with genotype ${x}_{i}$, we compute ${x}_{i}^{T}{\mathop{\beta }\limits^{ \sim }}_{r}^{(1)}\,,{x}_{i}^{T}{\mathop{\beta }\limits^{ \sim }}_{r}^{(2)}\,,\ldots ,\,{x}_{i}^{T}{\mathop{\beta }\limits^{ \sim }}_{r}^{(B)}$ to approximate their posterior distribution of genetic liability, generate a 90% credible interval ${\rm{C}}{\rm{I}}\text{-}{g}_{ir}$ (90% credible interval of genetic liability of i_{th} individual in r_{th} replication) with 5% and 95% quantile of the distribution and check whether their genetic liability is contained in the credible interval $I(\,{g}_{ir}\in {\rm{C}}{\rm{I}}\text{-}{g}_{ir})$. We compute the empirical coverage for each individual as the mean across the 100 simulation replicates \({{\rm{c}}{\rm{o}}{\rm{v}}{\rm{e}}{\rm{r}}{\rm{a}}{\rm{g}}

{\rm{e}}}_{i}=\frac{1}{100}{\sum }_{r=1}^{100}I(\,{g}_{ir}\in {\rm{C}}{\rm{I}}\text{-}{g}_{ir}).\,\)

Ethics declarations

All research carried out in this study conformed with the principles of the Helsinki Declaration. All individuals provided written informed consent to the original recruitment of the UCLA ATLAS Community Health Initiative. Patient Recruitment and Sample Collection for Precision Health Activities at UCLA is an approved study by the UCLA Institutional Review Board (Institutional Review Board number 17-001013). All analyses in this study use de-identified data (without any protected health information) with no possibility of re-identifying any of the participants.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The individual-level genotype and phenotype data of UKBB are available by application from http://www.ukbiobank.ac.uk/. Owing to privacy concerns, de-identified individual-level data for UCLA ATLAS are available only to UCLA researchers and can be accessed through the Discovery Data Repository Dashboard (https://it.uclahealth.org/about/ohia/ohia-products/discovery-data-repository-dashboard-0). Summary ATLAS association statistics are publicly available at https://atlas-phewas.mednet.ucla.edu/. PGSs for 84 traits investigated in this manuscript are available at the PGS Catalog (PGP000457, https://www.pgscatalog.org/publication/PGP000457). MCMC samplings of PGSs are available at https://doi.org/10.6084/m9.figshare.22413970.

Code availability

Scripts for simulations and real-data analyses are available at https://github.com/yidingdd/individual-pgs-accuracy (ref. [72]).

References

1. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

2. Mostafavi, H. et al. Variable prediction accuracy of polygenic scores within an ancestry group. *Elife* **9**, e48376 (2020).

3. Wang, Y., Tsuo, K., Kanai, M., Neale, B. M. & Martin, A. R. Challenges and opportunities for developing more generalizable polygenic risk scores. *Annu. Rev. Biomed. Data Sci.* **5**, 293–320 (2022).

4. Lambert, S. A., Abraham, G. & Inouye, M. Towards clinical utility of polygenic risk scores. *Hum. Mol. Genet.* **28**, R133–R142 (2019).

5. Johnson, R. et al. Leveraging genomic diversity for discovery in an electronic health record linked biobank: the UCLA ATLAS Community Health Initiative. *Genome Med.* **14**, 104 (2022).

6. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

7. Lewis, A. C. F. et al. Getting genetic ancestry right for science and society. *Science* **376**, 250–252 (2022).

8. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 581–590 (2018).

9. Kullo, I. J. et al. Polygenic scores in biomedical research. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-022-00470-z (2022).

10. Li, R., Chen, Y., Ritchie, M. D. & Moore, J. H. Electronic health records and polygenic risk scores for predicting disease risk. *Nat. Rev. Genet.* **21**, 493–502 (2020).

11. Natarajan, P. et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* **135**, 2091–2101 (2017).

12. Lee, A. et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet. Med.* **21**, 1708–1718 (2019).

13. Khera, A. V. et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* **177**, 587–596 (2019).

14. Perkins, D. O. et al. Polygenic risk score contribution to psychosis prediction in a target population of persons at clinical high risk. *Am. J. Psychiatry* **177**, 155–163 (2020).

15. Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat. Med.* **27**, 1876–1884 (2021).

16. Rodriguez, F. et al. Atherosclerotic cardiovascular disease risk prediction in disaggregated Asian and Hispanic subgroups using electronic health records. *J. Am. Heart Assoc.* **8**, e011874 (2019).

17. Jee, Y. H. et al. Validating breast cancer risk prediction models in the Korean Cancer Prevention Study-II Biobank. *Cancer Epidemiol. Biomarkers Prev.* **29**, 1271–1277 (2020).

18. Myers, P. D. et al. Identifying unreliable predictions in clinical risk models. *NPJ Digit. Med.* **3**, 8 (2020).

19. 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

20. Meuwissen, T. H., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829 (2001).

21. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).

22. Goddard, M. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* **136**, 245–257 (2009).

23. Goddard, M. E., Hayes, B. J. & Meuwissen, T. H. E. Using the genomic relationship matrix to predict the accuracy of genomic selection. *J. Anim. Breed. Genet.* **128**, 409–421 (2011).

24. Daetwyler, H. D., Villanueva, B. & Woolliams, J. A. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE* **3**, e3395 (2008).

25. Visscher, P. M., Yang, J. & Goddard, M. E. A commentary on "Common SNPs explain a large proportion of the heritability for human height" by Yang et al. (2010). *Twin Res. Hum. Genet.* **13**, 517–524 (2010).

26. Wray, N. R. et al. Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* **14**, 507–515 (2013).

27. Wientjes, Y. C. J. et al. Empirical and deterministic accuracies of across-population genomic prediction. *Genet. Sel. Evol.* **47**, 5 (2015).

28. Falconer, D. S. *Introduction to Quantitative Genetics* (Longman Scientific & Technical, 1989).

29. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* Vol. 1 (Sinauer Sunderland, 1998).

30. Walsh, B. & Lynch, M. *Evolution and Selection of Quantitative Traits* (Oxford Univ. Press, 2018).

31. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).

32. Ben Zaabza, H., Mäntysaari, E. A. & Strandén, I. Using Monte Carlo method to include polygenic effects in calculation of SNP-BLUP model

reliability. *J. Dairy Sci.* **103**, 5170–5182 (2020).

33. Pszczola, M., Strabel, T., Mulder, H. A. & Calus, M. P. L. Reliability of direct genomic values for animals with different relationships within and to the reference population. *J. Dairy Sci.* **95**, 389–400 (2012).

34. Wientjes, Y. C. J., Veerkamp, R. F. & Calus, M. P. L. The effect of linkage disequilibrium and family relationships on the reliability of genomic prediction. *Genetics* **193**, 621–631 (2013).

35. Habier, D., Fernando, R. L. & Dekkers, J. C. M. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* **177**, 2389–2397 (2007).

36. Lee, S. H., Weerasinghe, W. M. S. P., Wray, N. R., Goddard, M. E. & van der Werf, J. H. J. Using information of relatives in genomic prediction to apply effective stratified medicine. *Sci. Rep.* **7**, 42091 (2017).

37. Truong, B. et al. Efficient polygenic risk scores for biobank scale data by exploiting phenotypes from inferred relatives. *Nat. Commun.* **11**, 3074 (2020).

38. Martin, A. R. et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).

39. Privé, F. et al. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *Am. J. Hum. Genet.* **109**, 12–23 (2022).

40. Bitarello, B. D. & Mathieson, I. Polygenic scores for height in admixed populations. *G3* **10**, 4027–4036 (2020).

41. Wang, Y. et al. Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. *Nat. Commun.* **11**, 3865 (2020).

42. Coop, G. Genetic similarity versus genetic ancestry groups as sample descriptors in human genetics. Preprint at https://arxiv.org/abs/2207.11595 (2022).

43. Mathieson, I. & Scally, A. What is ancestry? *PLoS Genet.* **16**, e1008624 (2020).

44. Krainc, T. & Fuentes, A. Genetic ancestry in precision medicine is reshaping the race debate. *Proc. Natl Acad. Sci. USA* **119**, e2203033119 (2022).

45. Belbin, G. M. et al. Toward a fine-scale population health monitoring system. *Cell* **184**, 2068–2083 (2021).

46. Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).

47. Clarke, S. L. et al. Race and ethnicity stratification for polygenic risk score analyses may mask disparities in Hispanics. *Circulation* **146**, 265–267 (2022).

48. Johnson, R. et al. The UCLA ATLAS Community Health Initiative: promoting precision health research in a diverse biobank. *Cell Genomics* **3**, 100243 (2023).

49. Ding, Y. et al. Large uncertainty in individual polygenic risk score estimation impacts PRS-based risk stratification. *Nat. Genet.* **54**, 30–39 (2021).

50. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

51. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinform.* **12**, 246 (2011).

52. Gorjanc, G., Bijma, P. & Hickey, J. M. Reliability of pedigree-based and genomic evaluations in selected populations. *Genet. Sel. Evol.* **47**,

65 (2015).

53. Privé, F., Arbel, J. & Vilhjálmsson, B. J. LDpred2: better, faster, stronger. *Bioinformatics* **36**, 5424–5431 (2021).

54. Reich, D. et al. Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene. *PLoS Genet.* **5**, e1000360 (2009).

55. Scutari, M., Mackay, I. & Balding, D. Using genetic distance to infer the accuracy of genomic prediction. *PLoS Genet.* **12**, e1006288 (2016).

56. Fatumo, S. et al. A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).

57. Ruan, Y. et al. Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).

58. Spence, J. P., Sinnott-Armstrong, N., Assimes, T. L. & Pritchard, J. K. A flexible modeling and inference framework for estimating variant effect sizes from GWAS summary statistics. Preprint at *bioRxiv* https://doi.org/10.1101/2022.04.18.488696 (2022).

59. Zhang, H. et al. Novel methods for multi-ancestry polygenic prediction and their evaluations in 3.7 million individuals of diverse ancestry. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.24.485519 (2022).

60. Martin, A. R. et al. Increasing diversity in genomics requires investment in equitable partnerships and capacity building. *Nat. Genet.* **54**, 740–745 (2022).

61. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).

62. Hou, K. et al. Accurate estimation of SNP-heritability from biobank-scale data irrespective of genetic architecture. *Nat. Genet.* **51**, 1244–1251 (2019).

63. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).

64. Zhang, Q., Privé, F., Vilhjálmsson, B. & Speed, D. Improved genetic prediction of complex traits from individual-level data or summary statistics. *Nat. Commun*. **12**, 4192 (2021).

65. Lloyd-Jones, L. R. et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nat. Commun.* **10**, 5086 (2019).

66. Abraham, G. & Inouye, M. Fast principal component analysis of large-scale genome-wide data. *PLoS ONE* **9**, e93766 (2014).

67. Sorensen, D. & Gianola, D. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics* (Springer, 2002).

68. Behar, D. M. et al. No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. *Hum. Biol.* **85**, 859–900 (2013).

69. Abraham, G., Qiu, Y. & Inouye, M. FlashPCA2: principal component analysis of biobank-scale genotype datasets. *Bioinformatics* **33**, 2776–2778 (2017).

70. Privé, F., Aschard, H., Ziyatdinov, A. & Blum, M. G. B. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinformatics* **34**, 2781–2787 (2018).

71. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

72. Ding, Y. yidingdd/individual-pgs-accuracy. Zenodo https://doi.org/10.5281/zenodo.7787462 (2023)

Acknowledgements

We thank D. Geschwind, Y. Wu, R. Mester, R. Johnson and K. S. Burch for insightful comments. We gratefully acknowledge the resources provided by the Institute for Precision Health and participating patients from the UCLA ATLAS Community Health Initiative. The UCLA ATLAS Community Health Initiative in collaboration with UCLA ATLAS Precision Health Biobank is a programme of the Institute for Precision Health, which directs and supports the biobanking and genotyping of biospecimen samples from participating patients from UCLA in collaboration with the David Geffen School of Medicine, UCLA Clinical and Translational Science Institute and UCLA Health. The ATLAS Community Health Initiative is supported by UCLA Health, the David Geffen School of Medicine and a grant from the UCLA Clinical and Translational Science Institute (UL1TR001881). This research was conducted using the UKBB resource under application 33127. We thank the participants of UKBB for making this work possible. This work was financially supported in part by National Institutes of Health awards U01HG011715, R01HG009120 and R01MH115676. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author information

Authors and Affiliations

1. Bioinformatics Interdepartmental Program, UCLA, Los Angeles, CA, USA

   Yi Ding, Kangcheng Hou, Aditya Pimplaskar, Kristin Boulier & Bogdan Pasaniuc

2. Department of Computer Science, UCLA, Los Angeles, CA, USA

   Ziqi Xu & Ella Petter

3. National Centre for Register-based Research, Aarhus University, Aarhus, Denmark

   Florian Privé & Bjarni J. Vilhjálmsson

4. Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark

   Bjarni J. Vilhjálmsson

5. Novo Nordisk Foundation Center for Genomic Mechanisms of Disease, Broad Institute, Cambridge, MA, USA

   Bjarni J. Vilhjálmsson

6. Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

   Loes M. Olde Loohuis

7. Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

   Loes M. Olde Loohuis & Bogdan Pasaniuc

8. Department of Computational Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

   Bogdan Pasaniuc

9. Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

   Bogdan Pasaniuc

10. Institute for Precision Health, UCLA, Los Angeles, CA, USA

    Bogdan Pasaniuc

Contributions

Y.D. and B.P. conceived and designed the experiments. Y.D., K.H., Z.X., A.P., E.P. and K.B. carried out the experiments and statistical analyses. F.P.,

B.J.V. and L.M.O.L. provided statistical support. Y.D., K.H. and B.P. wrote the manuscript with the participation of all authors.

Corresponding authors

Correspondence to Yi Ding or Bogdan Pasaniuc.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 The individual level accuracy is highly correlated with population level accuracy.

All UKBB testing individuals are divided into 100 bins based on their GD. The x-axis is the average individual-level PGS accuracy for the individuals within the bin and the y-axis is (a) the squared correlation between simulated genetic liability and PGS estimates for the individuals within the bin (b) the squared correlation between simulated phenotype and PGS estimates. The

dot and error bars represent the mean and ±1.96 s.e.m from 100 simulations. Both p-values were derived from two-sided Pearson correlation tests without adjustment for multiple hypothesis testing. Any p-value below $\left({10}^{-10}\right)$ is annotated as $p < {10}^{-10}$.

Extended Data Fig. 2 PGS performance varies across GD in simulations using CB and NG as training data (${\bf{h}}_{\bf{g}}^{\bf{2}}{\boldsymbol{=}}{\bf{0.8}}$ and ${\bf{p}}_{\bf{causal}}={\bf{0.1}}{\boldsymbol{\%}}$).

(a) The coverage of the 90% credible intervals of genetic liability (CI-g_i) is approximately uniform across testing individuals at all GDs. The red dotted line represents the expected coverage of 90% CI-g_i. Each dot represents a randomly selected UKBB testing individual. For each dot, the x-axis is its GD from African training data, the y-axis is the empirical coverage of 90% CI-g_i calculated as the proportion of simulation replicates where the 90% credible intervals contain the individual’s true genetic liability, and the error bars represent mean ±1.96 standard error of the mean (s.e.m) of the empirical coverage calculated from 100 simulations. (b) The width of 90% CI-g_i increases with GD. For each dot, the y-axis is the average width of 90% CI-g_i across 100 simulation replicates, and the error bars represent ±1.96 s.e.m. (c) Individual PGS accuracy decreases with GD. For each dot, the y-axis is the average individual level PGS accuracy across 100 simulation replicates, and the error bars represent ±1.96 s.e.m. (d) Population-level metrics of PGS accuracy recapitulates the decay in PGS accuracy across genetic continuum. All UKBB testing individuals are divided into 100 equal-interval bins based on their GD. The x-axis is the average GD for the bin and the y-axis is the squared correlation between genetic liability and PGS estimates for the individuals within the bin. The dot and error bars represent the mean and ±1.96 s.e.m from 100 simulations.

Extended Data Fig. 3 PGS performance varies across GD in simulations using CB and NG as training data (${\bf{h}}_{\bf{g}}^{\bf{2}}{\boldsymbol{=}}{\bf{0.8}}$ and ${\bf{p}}_{\bf{causal}}={\bf{1}}{\boldsymbol{\%}}$).

(a) The coverage of the 90% credible intervals of genetic liability (CI-g_i) is approximately uniform across testing individuals at all GDs. (b) The width of 90% CI-g_i increases with GD. (c) Individual PGS accuracy decreases with GD. (d) Population-level metrics of PGS accuracy recapitulates the decay in PGS accuracy across genetic continuum. See Extended Data Fig. 2 for a detailed figure description.

Extended Data Fig. 4 The effect of different metrics of GD on the correlation between GD and accuracy.

The y-axis $-{cor}\left({r}_{i}^{2}\,,{d}_{i}\right)$ is the correlation between the GD and PGS accuracy; a larger correlation means GD has a better prediction of accuracy. The x-axis are different GD metrics: (1) GD based on PCA with varying number of PCs (from J = 1 to J = 20) and (2) GD based on GRM using pruned PCA SNPs only or all SNPs in PGS models. The GRM GD is computed as ${d}_{i}({\rm{GRM}})=\sqrt{\frac{1}{K}{\sum }_{k=1}^{K}{\left({x}_{i}-{x}_{k}\right)}^{2}}$, where ${x}_{i}$ is the standardized genotype of ${i}_{{th}}$ testing individual and ${x}_{k}$ is the standardized genotype of ${k}_{{th}}$ training individual.

Extended Data Fig. 5 The individual-level accuracy for height PGS decreases across the genetic ancestry continuum in UKBB.

(a) Individual PGS accuracy decreases within subcontinental GIA clusters. Each dot represents a testing individual from UKBB. For each dot, the x-axis represents its distance from the training population on the genetic continuum; the y-axis represents its PGS accuracy. The color represents the GIA cluster. (b) Individual PGS accuracy decreases across the entire UKBB. (c) The population PGS accuracy decreases with the average GD in each bin. All UKBB individuals are divided into 20 equal-interval GD bins. The x-axis is the average GD within the bin; the y-axis is the squared correlation between PGS and phenotype for individuals in the bin. The dot and error bar show mean and 95% confidence interval from 1000 bootstrap samples. *R* and *p* refer to the correlation between GD and PGS accuracy and its significance from two-sided Pearson correlation tests without adjustment for

multiple hypothesis testing. Any p-value below ${10}^{-10}$ is shown as $p < {10}^{-10}$.

Extended Data Fig. 6 Lower heterogeneity within the genetic ancestry group corresponds to a lower correlation between genetic distance and individual PGS accuracy.

(a) The distribution of correlations between PGS accuracy and GD for 84 traits in ATLAS. (b) The distribution of correlations between PGS accuracy and GD for 84 traits in UKBB. The x-axis is the homogeneity of the genetic ancestry clusters measured as the variance of GD within a genetic ancestry cluster; a larger ${\rm{var}}({{\rm{d}}}_{{\rm{i}}})$ indicates a larger variation of genetic background. Each boxplot contains 84 points corresponding to the correlation between PGS accuracy and GD within the group specified by x-axis for each of the 84 traits. The box shows the first, second and third quartile of the 84 correlations, and whiskers extend to the minimum and maximum estimates located within 1.5 × IQR from the first and third quartiles, respectively.

Extended Data Fig. 7 Empirical PGS accuracy decreases with genetic distance in UKBB averaged across 84 traits.

(a) Empirical PGS accuracy decreases across subcontinental GIA clusters. (b) Empirical PGS accuracy decreases across bins of GD. The x-axis is the average GD for all individuals within each GIA cluster or GD bin; the y-axis is the accuracy for each GIA cluster or GD bin. The dot and error bar show mean and $\pm$1.96 standard error of the mean across 84 traits.

Extended Data Fig. 8 Discordant directions of phenotype/PGS-distance correlations in UKBB.

The x axis is the correlation between phenotype and GD and the y axis is the correlation between PGS estimates and GD for all 48,586 testing individuals in UKBB. Numerical results are reported in Supplementary Table 4.

Extended Data Fig. 9 The correlation of PGS/phenotype with GD within each ancestry clusters in UKBB.

Only traits that exhibit significant correlation between GD and PGS/phenotype are shown as dots in the figure. The Ashkenazi cluster is not included because no significant correlations are observed.

Extended Data Fig. 10 Comparison of individual accuracy for height in UKBB.

(a) Accuracy computed from equation (1), with $va{r}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ set as fixed heritability; (b) Accuracy computed from equation (1), with $va{r}_{\beta }({x}_{i}^{{\rm{\top }}}\beta)$ estimated from Monte Carlo sampling from prior distribution of β. (c) Empirical accuracy estimated as the squared correlation between PGS and height for each genetic distance bin. All UKBB individuals are divided into 20 equal-interval GD bins. The x-axis is the average GD within the bin; the y-axis is the squared correlation between PGS and phenotype for individuals in the bin. The dot and error bar show mean and 95% confidence interval from 1000 bootstrap samples. Both (a) and (b) reflect the decreasing trend of empirical accuracy in (c). All p-values were derived from two-sided Pearson correlation tests without adjustment for multiple hypothesis testing. Any p-value below ${10}^{-10}$ is shown as $p < {10}^{-10}$.

Supplementary information

Supplementary Information

This file contains Supplementary Note, Figs. 1 and 2. Supplementary Note: Theoretical justification for individual PGS accuracy metric. Supplementary Fig. 1: Measured phenotype, PGS estimates, and accuracy varies across the EA GIA in ATLAS. Supplementary Fig. 2: Slope of regressing phenotype on PGS is calibrated across GIA clusters in simulation. See PDF for full legend.

Reporting Summary

Peer Review File

Supplementary Tables

This file contains Supplementary Tables 1–4. See Supplementary Information for legends.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Ding, Y., Hou, K., Xu, Z. *et al.* Polygenic scoring accuracy varies across the genetic ancestry continuum. *Nature* **618**, 774–781 (2023). https://doi.org/10.1038/s41586-023-06079-4

- Received: 28 September 2022
- Accepted: 12 April 2023
- Published: 17 May 2023

- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06079-4

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

A continuous measure for understanding the accuracy of genetically based predictions

Nature Research Briefing 17 May 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06079-4

- Article
- Open Access
- Published: 07 June 2023

The illusion of moral decline

- Adam M. Mastroianni &
- Daniel T. Gilbert

Nature **volume 618**, pages 782–789 (2023)

- 178k Accesses

- 2172 Altmetric

- Metrics details

Abstract

Anecdotal evidence indicates that people believe that morality is declining[1,2]. In a series of studies using both archival and original data (n = 12,492,983), we show that people in at least 60 nations around the world believe that morality is declining, that they have believed this for at least 70 years and that they attribute this decline both to the decreasing morality of individuals as they age and to the decreasing morality of successive generations. Next, we show that people's reports of the morality of their contemporaries have not declined over time, suggesting that the perception of moral decline is an illusion. Finally, we show how a simple mechanism based on two well-established psychological phenomena (biased exposure to information and biased memory for information) can produce an illusion of moral decline, and we report studies that confirm two of its predictions about the circumstances under which the perception of moral decline is attenuated, eliminated or reversed (that is, when respondents are asked about the morality of people they know well or people who lived before the

respondent was born). Together, our studies show that the perception of moral decline is pervasive, perdurable, unfounded and easily produced. This illusion has implications for research on the misallocation of scarce resources[3], the underuse of social support[4] and social influence[5].

Main

The social fabric appears to be unravelling: civility seems like an old-fashioned habit, honesty like an optional exercise and trust like the relic of another time. Some observers[6] claim that "the process of our moral decline" began with the "sinking of the foundations of morality" and proceeded to "the final collapse of the whole edifice", which brought us "finally to the dark dawning of our modern day, in which we can neither bear our immoralities nor face the remedies needed to cure them". But as apt as this description of our times may seem, it was written more than 2,000 years ago by the historian Livy, who was bemoaning the declining morality of his fellow Roman citizens. From ancient to modern times, social observers have often lamented the ugly turns their societies have taken, and have often suggested that a recent decline in morality—in kindness, honesty and basic human decency—was among the causes[2,7].

Why have so many different people in so many different times and places been convinced that their fellow citizens are now less moral than they once were? One possibility is that morality has, in fact, been declining worldwide for millennia—declining so steadily and so precipitously that people in every era have been able to observe that decline in the brief span of a human lifetime. The other possibility is that the perception of moral decline is a psychological illusion to which people all over the world and throughout history have been susceptible. We provide evidence for the latter possibility. First, we show that people in at least 60 nations do indeed believe that morality is declining, and that they have believed this for at least 70 years. Second, we show that people attribute this decline both to the decreasing morality of individuals as they age and to the decreasing morality of successive generations. Third, we show that people's reports of the current morality of their contemporaries have not declined over time, which strongly suggests that the perception of moral decline is an illusion. Fourth and finally, we describe tests of a simple psychological mechanism that can

produce the illusion of moral decline and can predict some of the circumstances under which it will be attenuated, eliminated or reversed (for example, when respondents are asked about the morality of people they know well or people who lived before the respondent was born).

Do people perceive moral decline?

Morality refers primarily to people's treatment of each other[8], which ranges from the altruistic[9] to the barbaric[10]. But like most social observers, Livy was not remarking on the moral extremes—on the rare heroic deed or occasional heinous crime that few people ever perform or experience. Rather, he was remarking on the ways in which ordinary people behave in their daily lives. Do modern people, like Livy, believe that their contemporaries are less honest and kind than they used to be? Do they think their neighbours are less generous and less helpful, that their co-workers are more likely to treat each other disrespectfully and betray each other's trust? Survey researchers have been asking people about their perceptions of changes in these everyday moral qualities since at least 1949, but the full corpus of relevant survey data has never been systematically assembled and analysed. We began by doing that.

In study 1, we searched the databases of major survey research providers (using search terms shown in the Supplementary Information) and found 177 survey items that asked representative samples of a total of 220,772 US Americans if and how they thought other people's morality had changed over time (Supplementary Table 1). These items were administered over a 70-year span from 1949 to 2019. Typical items included: "Do you think that over the last few decades our society has become less honest and ethical in its behavior, more honest and ethical, or has there been no change in the extent to which people behave honestly and ethically?" and "Right now, do you think the state of moral values in this country as a whole is getting better or getting worse?" (further methodological details can be found in the concluding method section as well as in the Supplementary Information). On 84.18% of the items, the majority of participants reported that morality had declined. A linear model indicated that the proportion of participants who reported moral decline was not significantly influenced by the year in which the survey was administered, $b = 0.07$, 95% confidence interval (CI) =

[−0.11, 0.24], $t(175) = 0.77$, $P = 0.45$, adjusted $R^2 = -0.002$, and the same model fit in a Bayesian framework indicated strong evidence of no effect (Bayes Factor of 0.04), which is to say that US Americans have been reporting moral decline at the same rate for as long as researchers have been asking them about it. (These and all tests we report are two-tailed).

Two more findings were noteworthy. First, participants in study 1 were more likely to perceive moral decline when they were asked about longer periods of time (for example, “the last decade”) than about shorter periods of time (“the last year”), $b = 0.57$, 95% CI = [0.09, 1.05], $t(43) = 2.42$, $P = 0.02$, adjusted $R^2 = 0.10$, which is precisely what one would expect if participants believed that morality has been declining continuously. Second, participants reported increases in morality when asked about a few specific issues on which social progress has clearly been made: for example, 59% of participants reported improved treatment of African Americans, 51% reported improved treatment of people with physical disabilities and 50% reported improved treatment of gay people. The fact that participants calculated moral decline cumulatively across time periods and acknowledged special exceptions to the general trend suggests that they were reporting well-considered beliefs, and not merely expressing some vague sense of despair about humanity. Indeed, in the Supplementary Information, we report an extra study (Supplementary study 3) showing that the perception of moral decline persists even when people are incentivized to respond accurately.

The perception of moral decline was not unique to US Americans. We resampled the databases of major survey research providers and found 58 survey items that asked a total of 354,120 participants in 59 nations other than the United States if and how they thought other people’s morality had changed over time (Supplementary Table 2). These items were administered over a 13-year span from 1996 to 2007. An analysis of these items showed that on 86.21% of the items, the majority of non-US participants reported that morality had declined. Indeed, the Pew Research Center surveyed citizens of 40 nations in 2002 (ref. [11]) and 2006 (ref. [12]) and, as Fig. 1 shows, in every one of those nations, the majority of participants reported that moral decline was at least a “moderately big problem”.

Fig. 1: Countries surveyed by Pew in 2002 or 2006.



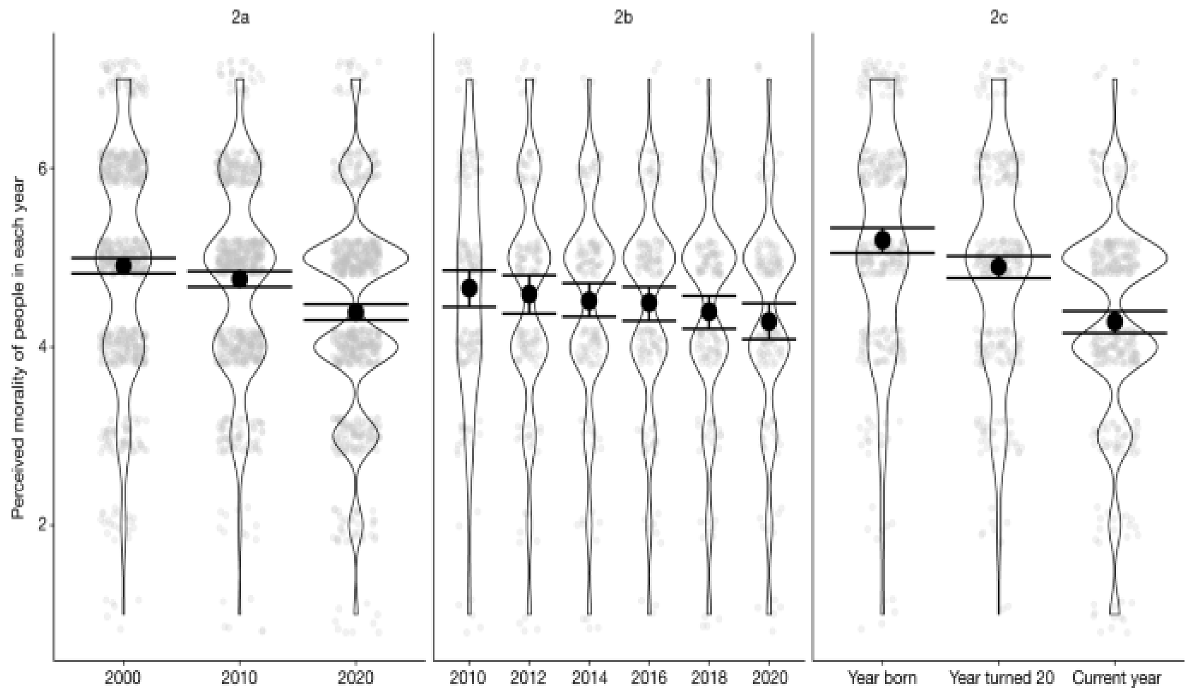
In every country surveyed by Pew in 2002 or 2006 (shown in red), the majority of participants reported that moral decline was at least a “moderately big problem”. Map created with MapChart.

The survey items we analysed in study 1 (Supplementary Tables 1 and 2) used a wide range of question formats to ask participants over a wide range of decades about moral decline across a wide range of time periods, and they converged on a single conclusion: people all over the world believe that morality has declined, and they have believed this for as long as researchers have been asking them about it. Archival data are uniquely able to tell us how people in the past thought and felt, but they have limits. Some of the items we analysed asked participants for their perceptions of changes in “moral values” without specifying what those values were (for example, “Right now, do you think the state of moral values in this country as a whole is getting better or getting worse?”), some failed to specify the time in the past to which the present was to be compared (for example, “Compared to the past, are people today more or less friendly toward their neighbors?”) and some contained ambiguous wording that was not optimal for extracting accurate measures of people’s perceptions of moral decline (“Considering just the moral climate of the country today, do you feel things in this country are generally going in the right direction or do you feel things have pretty seriously gotten off on the wrong track?”). In addition, all items asked

participants questions about the presence or absence of moral decline rather than asking them to rate the level of morality of people in both the present and the past, which allowed us to compute the proportion of participants who perceived moral decline but not how much decline they perceived. We addressed these and other limitations of the archival data by conducting three original studies.

In studies 2a–c, we asked samples of US Americans to rate how "kind, honest, nice, and good" people were in 2020 (the year the studies were conducted), as well as in various other years that differed by study. Methodological details can be found in the concluding method section and in the Supplementary Information. As Fig. 2 shows, participants in study 2a ($n = 698$ respondents on Prolific), study 2b ($n = 185$ respondents on Amazon Mechanical Turk) and study 2c ($n = 347$ respondents on Amazon Mechanical Turk) perceived moral decline. Specifically, in study 2a, participants rated people as less kind, honest, nice and good in 2020 (mean (M) = 4.39) than in 2010 ($M = 4.76$, $b = -0.37$, 95% CI = [−0.46, −0.28], $t(1394) = -9.38$, $P < 0.001$, Cohen's $d = -0.50$), or in 2000 ($M = 4.91$, $b = -0.52$, 95% CI = [−0.62, −0.43], $t(1394) = -13.23$, $P < 0.001$, $d = -0.71$), and as less kind, honest, nice and good in 2010 than in 2000, $b = -0.15$, 95% CI = [−0.25, −0.06], $t(1394) = -3.85$, $P < 0.001$, $d = -0.21$. We also conducted two direct replications of study 2a (one of which was preregistered; Supplementary Information), which produced the same results.

Fig. 2: Results of studies 2a–c.



The panels show the results of studies 2a (left panel), 2b (middle panel) and 2c (right panel). Opaque points represent means. Transparent points represent individual observations jittered for legibility. Error bars represent 95% CIs. Study 2a $n = 698$, study 2b $n = 148$ and study 2c $n = 347$.

In study 2b, participants rated people as less kind, honest, nice and good in 2020 ($M = 4.28$) than in 2016 ($M = 4.49$, $b = -0.21$, 95% CI = [−0.43, 0.006], $t(735) = -2.87$, $P = 0.047$, $d = -0.33$), in 2014 ($M = 4.51$, $b = -0.23$, 95% CI = [−0.45, −0.01], $t(735) = -3.14$; $P = 0.02$, $d = -0.37$), in 2012: $M = 4.59$, $b = -0.30$, 95% CI = [−0.52, −0.09], $t(735) = -4.16$, $P < 0.001$, $d = -0.48$) and in 2010 $M = 4.66$, $b = -0.37$, 95% CI = [−0.59, −0.16], $t(735) = -5.08$, $P < 0.001$, $d = -0.59$). Participants in study 2b also rated people in 2018 as less kind, honest, nice and good than in people in 2010 ($b = -0.26$, 95% CI = [−0.48, −0.05], $t(735) = -3.61$, $P = 0.004$, $d = -0.32$). No other comparisons were significant (all $P > 0.05$).

In study 2c, participants rated people as less kind, honest, nice and good in 2020 ($M = 4.28$) than in the year the participant turned 20 years old ($M = 4.89$, $b = -0.61$, 95% CI = [−0.77, −0.45], $t(675) = -9.21$, $P < 0.001$, $d = -0.72$) and in the year the participant was born ($M = 5.20$, $b = -0.92$, 95% CI = [−1.07, −0.76], $t(667) = -14.27$, $P < 0.001$, $d = -1.08$). Participants also rated people as less kind, honest, nice and good in the year the participant

turned 20 years old than in the year the participant was born, $b = -0.31$, 95% CI = [−0.47, −0.15], $t(675) = -4.70$, $P < 0.001$, $d = -0.37$.

In studies 2a–c, we also examined the effects of age, gender, race, education (1 = did not finish high school; 6 = graduate degree), political ideology (−2 = very liberal; 2 = very conservative) and parental status (0 = not parent; 1 = parent) on perceptions of moral decline using an exploratory linear regression. In study 2a, more conservative participants perceived more decline, $b = -0.18$, 95% CI = [−0.27, −0.09], $t(684) = -4.03$, $P < 0.001$. No other effects in study 2a were significant (all $P > 0.05$). An additional exploratory one-sample t-test indicated that, although more conservative participants perceived more moral decline than did more liberal participants, more liberal participants perceived moral decline as well, $M = -0.35$, 95% CI = [−0.47, −0.23], $t(398) = -4.96$, $d = 0.28$, $P < 0.001$, $d = 0.28$. In study 2b, no other effects were significant (all $P > 0.05$). In study 2c, more conservative participants perceived more moral decline than did more liberal participants, $b = -0.15$, 95% CI = [−0.28, −0.12], $t(329) = -2.15$, $P = 0.03$, but more liberal participants perceived moral decline as well, $M = -0.80$, 95% CI = [−1.02, −0.58], $t(159) = -7.23$, $P < 0.001$, $d = 0.57$.

In study 2c, older participants perceived more moral decline than did younger participants, $b = -0.02$, 95% CI = [−0.03, −0.004], $t(329) = -2.61$, $P = 0.01$. No other effects were significant (all $P > 0.05$). We further investigated the effect of age in study 2c by creating two moral decline scores: specifically, (1) we subtracted participants' ratings of people in the year the participant was 20 years old from their ratings of people in 2020 and (2) we subtracted participants' ratings of people in the year the participant was born from their ratings of people in 2020. An exploratory linear model indicated that older participants perceived more moral decline than did younger participants both compared to the year in which they turned 20 years old: $b = -0.02$, $t(320) = -3.30$, 95% CI = [−0.03, −0.007], $P < 0.001$, and compared to the year in which they were born: $b = -0.02$, $t(345) = -3.77$, 95% CI = [−0.04, −0.01], $P < 0.001$. Did older participants perceive more moral decline simply because they were considering longer periods of time? Yes. We created a measure of the annual rate of moral decline by subtracting the participants' ratings of people in the year the participant was born from their rating of people in 2020, and then dividing that value by the participant's age. We then fit an exploratory linear model with the perceived

annual rate of moral decline as the outcome and age as a predictor. The main effect of age was not significant, $b = -0.0002$, 95% CI = [−0.0005, 0.0002], $t(345) = -1.05$, $P = 0.29$, and refitting the same model in a Bayesian framework provided strong evidence that perceived moral decline per year did not differ by age (100% of high-density interval (HDI) in the region of practical equivalence (ROPE)). In other words, younger and older participants did not report different annual rates of moral decline, which is to say that they reported different total amounts of moral decline only because they were reporting on moral decline across different numbers of years.

To what do people attribute moral decline?

People clearly perceive moral decline, but to what do they attribute it? There are two possibilities. The average morality of a population may decline between two points in time (T_1 and T_2) because (1) individuals who are moral at T_1 are less moral when they reach T_2 (a phenomenon we refer to as ‘personal change’), and/or (2) older people who were alive at T_1 but who died before T_2 are more moral than younger people who were alive at T_2 but who were not yet born (or who were not yet adults and therefore not sampled) at T_1 (a phenomenon we refer to as ‘interpersonal replacement’). When the average morality of a population declines over very short time periods (for example, a day), the decline is probably the result of personal change (because very few people who were measured at T_1 were not also measured at T_2), and when the average morality of a population declines over very long time periods (for example, 200 years), the decline is necessarily the result of interpersonal replacement (because no human being lives for 200 years and therefore no one who was measured at T_1 was also measured at T_2).

So, what about moral decline over the intermediate time periods that participants in studies 2a–c were asked about? To which of these sources—personal change or interpersonal replacement—do people attribute such decline? In study 3 ($n = 319$ respondents on Amazon Mechanical Turk), we asked a sample of US Americans to rate how “kind, honest, nice, and good” people were in 2020 (the year the study was conducted) and in 2005.

Methodological details can be found in the Methods and in the Supplementary Information. Next, participants rated the morality of two exclusive subsets of this population. The first subset was people who were living adults in both 2005 and 2020. The difference between these ratings was a measure of participants' perceptions of personal change between the 2 years. The second subset was people who were living adults in either 2005 or 2020, but not in both years. The difference between these ratings was a measure of participants' perceptions of interpersonal replacement between the 2 years.

A paired samples *t*-test indicated that participants perceived moral decline, rating people as less kind, honest, nice and good in 2020 ($M = 4.35$) than in 2005 ($M = 4.89$), $t(318) = -9.88$, 95% CI = [−0.65, −0.44], $d = 0.55$, $P < 0.001$. To determine whether participants attributed this moral decline to personal change and/or to interpersonal replacement, we used linear regression to determine whether and how well participants' perceptions of personal change and of interpersonal replacement predicted their perceptions of moral decline. Both measures significantly predicted participants' perceptions of moral decline (personal change $b = 0.50$, 95% CI = [0.40, 0.59], $t(316) = 9.96$, $P < 0.001$; interpersonal replacement $b = 0.17$, $t(316) = 6.52$, 95% CI = [0.12, 0.22], $P < 0.001$; adjusted $R^2 = 0.36$). We refit the model to include age, gender, race, political ideology, education and parental status as covariates, and the effects of personal change and interpersonal replacement both remained significant (personal change $b = 0.50$, 95% CI = [0.40, 0.60], $t(304) = 9.87$, $P < 0.001$; interpersonal replacement $b = 0.18$, 95% CI = [0.13, 0.23], $t(304) = 6.71$, $P < 0.001$; adjusted $R^2 = 0.38$).

In short, participants in study 3 believed that morality had declined on average over a 15-year period, and they attributed that decline both to the decreasing morality of individuals over time and to the decreasing morality of successive generations[13,14]. The fact that people attribute moral decline to both sources may help explain why their perceptions of moral decline are so robust, appearing in study 3, in the archival data of study 1 and in the original data collected for studies 2a–c.

Is morality declining?

People believe that morality is declining. Is it? Societies keep (or at least leave) reasonably good records of extremely immoral behaviour such as slaughter and conquest, slavery and subjugation or murder and rape, and careful analyses of those historical records strongly suggest that these objective indicators of immorality have decreased significantly over the last few centuries[15,16]. On average, modern humans treat each other far better than their forebears ever did—which is not what one would expect if honesty, kindness, niceness and goodness had been decreasing steadily, year after year, for millennia. Although there are no similarly objective historical records of everyday morality—of how often people offer their seats to an elderly person, give directions to a lost tourist or help their neighbour fix a fence—there are subjective measures of such things.

Recall that in study 1, we examined people's reports of moral change, which were obtained when survey researchers asked people to mentally compare the morality of people in the present to the morality of people at some point in the past and then report the direction of the difference. But, for decades, survey researchers have also been asking people to report directly on the moral values, traits and behaviours of themselves and their contemporaries in the present: "Were you treated with respect all day yesterday?" or "Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?" or "During the past 12 months, how often have you carried a stranger's belongings, like groceries, a suitcase, or shopping bag?" (Supplementary Tables 3 and 4). If, as people all over the world claim, morality has been declining steadily and precipitously for decades, then people's reports of current morality should also have declined over the years. Have they?

In study 4, we searched the databases of major survey research providers (using search terms listed in the Supplementary Information) and found 107 items that were administered to 4,483,136 people across a 55-year span from 1965 to 2020, and that (1) asked participants to report on some aspect of current morality and (2) were administered at least twice, at times that were at least 10 years apart (Supplementary Table 3 shows the items). To determine whether people's reports of the current morality of their contemporaries changed over time, we fit a linear model for each survey. The year of each survey was always entered as a predictor, and the outcome was always the average perception of current morality. Because these

surveys generally had large samples—some with hundreds of thousands of participants—the significance of *P* values is not very meaningful, so we used R^2 values as a measure of effect size. To shed further light on the size of these effects, we also fit analogous models in a Bayesian framework.

The results of both analyses were clear: people's reports of the current morality of their contemporaries were stable over time. On average, the year in which the survey was conducted explained less than 0.3% of the variance in responses, and in almost all cases it explained less than 1% (Supplementary Table 4). This result was confirmed by Bayesian analysis, which showed that 100% of the HDI was within the ROPE in all but one case, indicating that any changes over time were negligible at best. We repeated these analyses for data collected from non-US samples (33 samples, $n = 7{,}432{,}736$) and found similar results: on average, the year in which the survey was conducted explained less than 0.2% of the variance in responses (all items and results for the non-US sample are listed in Supplementary Table 4). In short, studies 1–3 showed that when people are explicitly asked to assess moral change, they claim that morality has declined, but study 4 shows that when people are asked to assess the current morality of their contemporaries, their assessments do not change over time.

Could this be because words can change meaning over time? If residents of Los Angeles in both 1942 and 2022 described traffic as 'heavy', it would be a mistake to conclude that traffic had not actually increased. Words such as 'heavy' and 'moral' are inherently ambiguous, and if people adapt to changes in traffic or morality, then people in different decades may use the same ambiguous word to describe very different states of affairs. This is unlikely to be the case in study 4 because in addition to including a few items that measured traits and values with ambiguous terms such as 'morality', the dataset (Supplementary Tables 3 and 4) mainly contained items that measured specific and relatively unambiguous moral behaviours, such as "Within the past 12 months, have you been assaulted or mugged?" or "During the past 12 months, have you let a stranger go ahead of you in line?" Answers to specific and unambiguous questions such as these did not change over time. It seems rather improbable that people were less likely to allow strangers into a line in 2020 than in 2010, but that somehow in that 10-year span, the meaning of words such as 'stranger' and 'line' had changed in ways that masked that objective decline in kindness.

The subjective measures we analysed are not definitive, of course, but they strongly suggest that the widespread perception of moral decline is an illusion. Moreover, studies that use the rare objective measure of changes in everyday moral behaviour suggest the same thing. For instance, Yuan et al.[17] showed that rates of cooperation in the Prisoner's Dilemma game have increased significantly between 1956 and 2017, and in the Supplementary Information, we report the results of a study (Supplementary study 3) showing that most people mistakenly believe that such cooperation has declined.

Why do people perceive moral decline?

The results of studies 1–3 suggest that people believe that morality has declined, and the results of study 4 suggest that this belief is illusory. If morality has not declined, then why do people think it has? Although there are surely many good answers to this question, we suggest that one of them has to do with the fact that when two well-established psychological phenomena work in tandem, they can produce an illusion of moral decline. First, numerous studies have shown that human beings are especially likely to seek and attend to negative information about others[18,19,20], and mass media indulge this tendency with a disproportionate focus on people behaving badly[21]. As such, people may encounter more negative information than positive information about the morality of 'people in general', and this 'biased exposure effect' may help explain why people believe that current morality is relatively low. Second, numerous studies have shown that when people recall positive and negative events from the past, the negative events are more likely to be forgotten[22], more likely to be misremembered as their opposite[23,24] and more likely to have lost their emotional impact[25]. This 'biased memory effect' may help explain why people believe that past morality was relatively high. Working together, these two phenomena can produce an illusion of moral decline. Specifically, biased exposure to information about current morality may make the present seem like a moral wasteland, biased memory for information about past morality may make the past seem like a moral wonderland and when people in a wasteland remember being in a wonderland, they may naturally conclude that the landscape has changed.

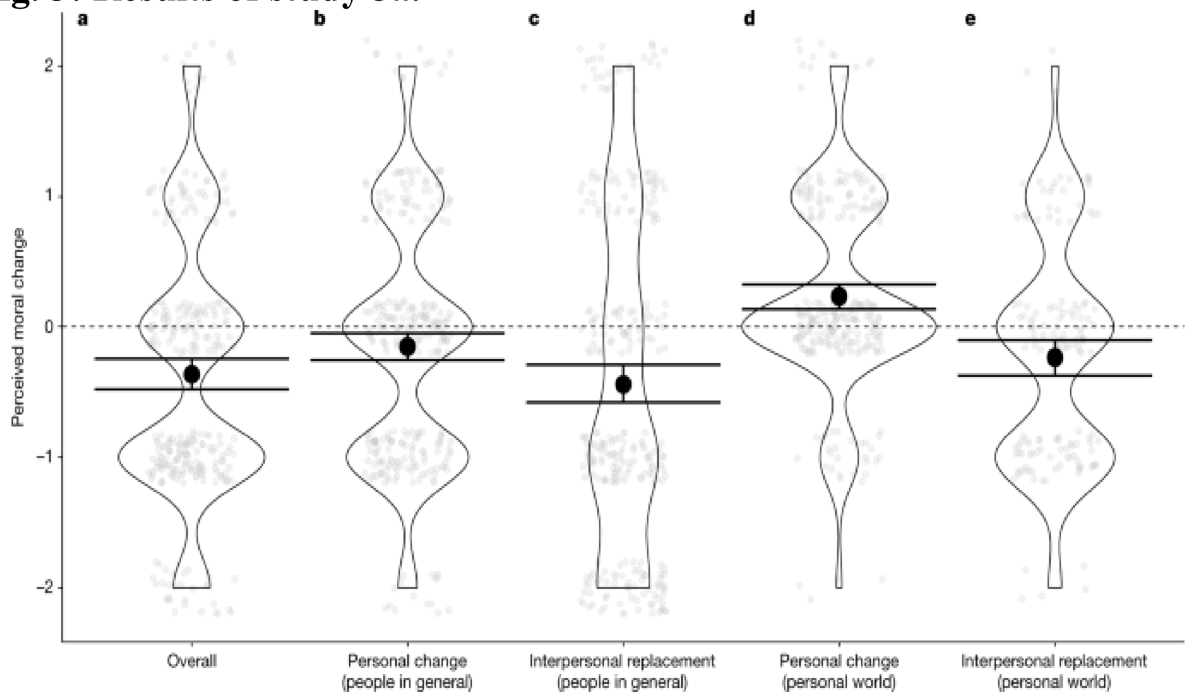
This 'biased exposure and memory' (BEAM) mechanism comports well with the results of the studies we have described, but it also makes at least two testable predictions. Specifically, the BEAM mechanism predicts that the illusion of moral decline should be attenuated, eliminated or even reversed when (1) people are exposed to a disproportionate amount of positive rather than negative information about the moral behaviour of others, as they are with their families, friends and associates, and (2) when people are asked about times for which they have little or no information in memory, such as in the years before they were born. In the Supplementary Information, we provide a mathematical model of the BEAM mechanism and show how the model makes these two predictions, which we tested in studies 5a and 5b.

Study 5a ($n = 283$ respondents on Amazon Mechanical Turk) tested the hypothesis that the illusion of moral decline is attenuated, eliminated or reversed when participants are asked to rate people in their personal worlds rather than people in general. As described in the Methods (and Supplementary Information), we began by measuring participants' perceptions of (1) overall moral decline; (2) personal change among people in general and (3) interpersonal replacement among people in general. Then we measured participants' perceptions of (4) personal change among people in their personal worlds and (5) interpersonal replacement among people in their personal worlds. We explained that the phrase 'personal worlds' referred to "all the people with whom you currently interact, in person or otherwise, in your everyday life. This probably includes friends, family members, coworkers, classmates, neighbors, etc.".

We used one-sample *t*-tests to determine whether each of the measures described above differed significantly from zero. First, participants on average perceived moral decline: they believed that people in general were not as kind, honest, nice and good in 2020 as they were in 2005 ($M = -0.36$), $t(282) = -6.04$, 95% CI = [−0.48, −0.25], $d = 0.36$, $P < 0.001$. Second, participants believed that individuals in 2020 were not as kind, honest, nice and good as those same individuals had been in 2005, $M = -0.15$, $t(282) = -2.67$, 95% CI = [−0.26, −0.04], $d = 0.16$, $P = 0.008$, and that younger people in 2020 were not as kind, honest, nice and good as older people were in 2005, $M = -0.44$, $t(282) = -5.82$, 95% CI = [−0.59, −0.29], $d = 0.35$, $P < 0.001$. In other words, as in study 3, participants believed that morality had

declined among individuals and between successive generations. These results are illustrated in Fig. 3.

Fig. 3: Results of study 5a.



a–**e**, From left to right, this figure shows the perceived difference in morality of (**a**) people in general in 2005 and people in general in 2020 (overall) (**b**), people in general who were sampled both in 2005 and 2020 (personal change among people in general) (**c**), people in general who were sampled in 2005 or 2020 but not in both years (interpersonal replacement among people in general) (**d**), people in the participant's personal world who were sampled both in 2005 and 2020 (personal change among people in personal world) and (**e**) people in the participant's personal world who were sampled in 2005 or 2020 but not in both years (interpersonal replacement among people in participant's personal world). Opaque points represent means. Transparent points represent individual observations jittered for legibility. Error bars represent 95% CIs. $n = 283$.

Did participants believe the same things about people in their personal worlds? No. First, participants believed that the individuals who were in their personal worlds in both 2005 and 2020 had shown moral improvement over that period rather than moral decline, $M = 0.23$, $t(255) = 4.86$, 95% CI = [0.14, 0.33], $d = 0.30$, $P < 0.001$. Second, although participants believed that

the younger people who were in their personal worlds in 2020 (but not in 2005) were not as kind, honest, nice and good as the older people who were in their personal worlds in 2005 (but not in 2020), $M = -0.23$, $t(144) = -3.23$, 95% CI = [−0.38, −0.09], $d = 0.27$, $P = 0.002$, this difference was smaller among people in their personal worlds than it was among people in general, $t(144) = -2.56$, 95% CI = [−0.40, −0.05], $d = 0.18$, $P = 0.01$.

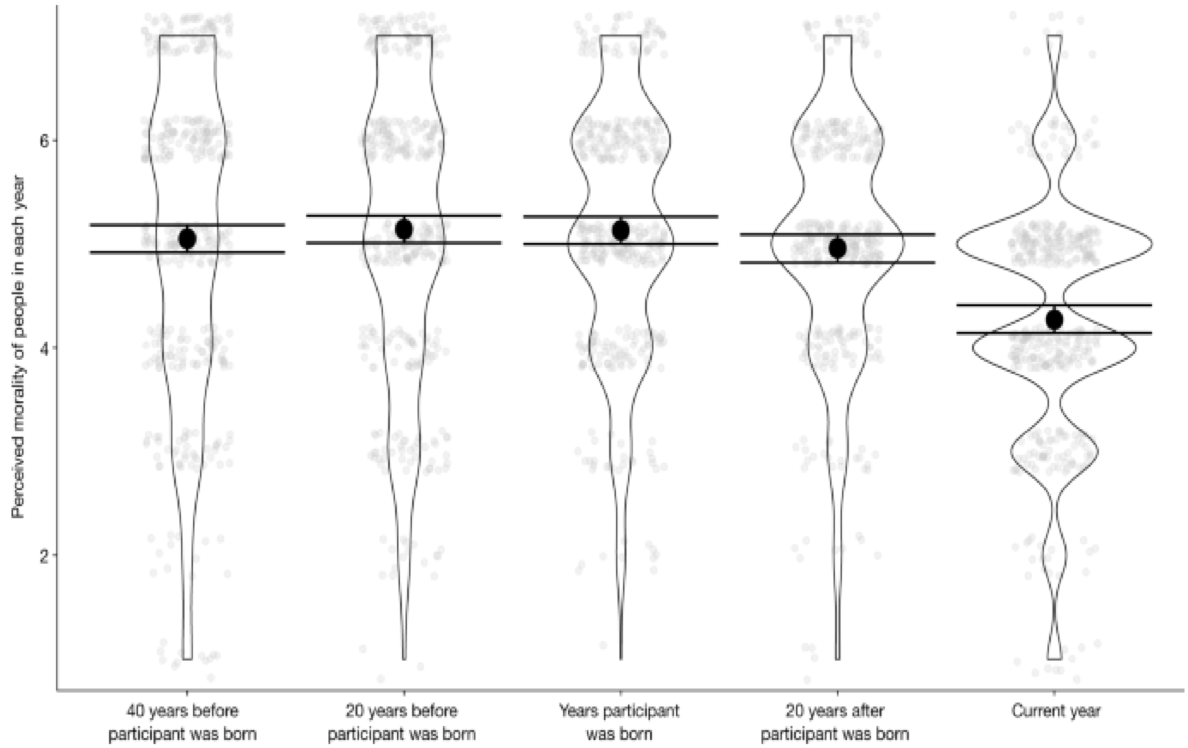
To investigate the effects of demographic variables on the perception of moral decline, we fit the same exploratory model used in studies 2a–c. The outcome variable was participants' perceptions of moral decline between 2005 and 2020. Older participants perceived more moral decline than did younger participants, $b = -0.02$, 95% CI = [−0.03, −0.005], $t(272) = -2.93$, $P = 0.004$, and non-parents perceived more moral decline than did parents, $b = 0.29$, 95% CI = [0.02, 0.56], $t(271) = 2.12$, $P = 0.03$, adjusted $R^2 = 0.07$. No other effects were significant (all $P > 0.05$).

In short, participants in study 5a believed that morality had declined among people in general, but this effect was reversed (in the case of personal change) or attenuated (in the case of interpersonal replacement) among the people they personally knew. We hasten to note that there are surely many reasons why people might think differently about people in their personal worlds than about people in general and that the BEAM mechanism is, at best, just one.

Study 5b ($n = 387$ respondents on Amazon Mechanical Turk) tested the hypothesis that the illusion of moral decline is attenuated, eliminated or reversed when participants are asked to rate the morality of people in general in the years before the participant was born. Participants rated how kind, honest, nice and good people in general are or were at four points in time: in the current year (which was 2021), 20 years after the participant was born, the year the participant was born, 20 years before the participant was born and 40 years before the participant was born. We fit the same model and planned contrasts used in studies 2a–c. As in our previous studies, participants perceived moral decline among people in general in the years after the participant was born. Specifically, participants believed that people in general were (1) less kind, honest, nice and good in 2021 ($M = 4.27$) than they were in the year the participant was 20 years old ($M = 4.96$), $b = –0.68$, 95% CI = [–0.85, –0.51], $t(1513) = –10.07$, $P < 0.001$, $d = -0.75$ and (2) less

kind, honest, nice and good in the year the participant was 20 years old than they were in the year the participant was born ($M = 5.13$), $b = –0.18$, 95% CI [–0.35, –0.01], $t(1513) = –2.60$, $P = 0.03$, $d = −0.19$. However, there was no evidence to suggest that participants perceived moral decline in the years before they were born. Specifically, there was no evidence that participants believed that people in general were (1) any more or less kind, honest, nice and good in the year the participant was born ($M = 5.13$) than they were 20 years before the participant was born ($M = 5.14$), $b = –0.01$, 95% CI = [–0.17, 0.15], $t(1506) = –0.16$, $P = 0.87$, $d = −0.01$ and (2) any more or less kind, honest, nice and good 20 years before the participant was born ($M = 5.14$) than they were 40 years before the participant was born ($M = 5.05$), $b = 0.09$, 95% CI = [−0.24, 0.04], $t(1506) = −1.42$, $P = 0.31$, $d = 0.10$. Equivalence tests using the ‘parameters’ package in R[26] indicated that there was insufficient evidence to conclude that participants’ ratings for 40 years before their birth and 20 years before their birth were equivalent (91.74% of HDI in ROPE, $P = 0.09$; if anything, participants perceived moral improvement between these years), but that there was sufficient evidence to conclude that participants’ ratings for 20 years before their birth and the year of their birth were equivalent (100% of HDI in ROPE, $P = 0.003$). In short, participants believed that moral decline began at about roughly the same time they appeared on Earth. These results are illustrated in Fig. 4.

Fig. 4: Results of Study 5b.



The figure shows the perceived morality of people in various years. Opaque points represent means. Transparent points represent individual observations jittered for legibility. Error bars represent 95% CIs. $n = 387$.

To investigate the effects of demographic variables on the perception of moral decline, we fit the same exploratory model used in studies 2a–c. The outcome variable was perceived moral decline between the year of the participant's birth and 2021. More conservative participants perceived more moral decline than did more liberal participants, $b = -0.33$, 95% CI = [−0.45, −0.22], $t(374) = -5.81$, $P < 0.001$, but a one-sample t-test indicated that more liberal participants perceived moral decline as well, $b = -0.46$, 95% CI = [−0.64, −0.28], $t(196) = -5.11$, $P < 0.001$, $d = 0.36$. Although older participants perceived more moral decline than did younger participants, $b = -0.01$, 95% CI = [−0.02, −0.003], $t(374) = -2.60$, $P = 0.009$, this was because older participants were perceiving moral decline over a longer period of time. Indeed, the same analysis used in study 2c indicated no evidence that older and younger participants perceived different annual rates of moral decline between the year they were born and 2021, $b = 0.000005$, 95% CI = [−0.0003, 0.0003], $t(385) = 0.04$, $P = 0.97$, consistent with the results of study 2c.

Studies 5a and 5b show that when participants were asked to assess the morality of people about whom they had mainly positive information in memory (that is, people in their personal worlds) or about whom they had little or no information in memory (that is, people who lived before the participants were born), the perception of moral decline was attenuated, eliminated or reversed, just as the BEAM mechanism predicts. The illusion of moral decline is a robust phenomenon that surely has several causes, and no one can say which of them produced the illusion that our studies have documented. Studies 5a and 5b do not directly implicate the BEAM mechanism in that production but they do make it a viable candidate for future research.

Discussion

Participants in the foregoing studies believed that morality has declined, and they believed this in every decade and in every nation we studied. They believed the decline began somewhere around the time they were born, regardless of when that was, and they believed it continues to this day. They believed the decline was a result both of individuals becoming less moral as they move through time and of the replacement of more moral people by less moral people. And they believed that the people they personally know and the people who lived before they did are exceptions to this rule. About all these things, they were almost certainly mistaken. One reason they may have held these mistaken beliefs is that they may typically have encountered more negative than positive information about the morality of contemporaries whom they did not personally know, and the negative information may have faded more quickly from memory or lost its emotional impact more quickly than the positive information did, leading them to believe that people today are not as kind, nice, honest or good as once upon a time they were.

Like all studies, ours have limitations. For example, studies 1 and 4 made use of archival data that were not collected for the purposes to which we put them and that were therefore less than ideal. For example, some of the items we analysed asked participants for their perceptions of changes in ‘moral values’ without specifying what those values were, some failed to specify the time in the past to which the present was to be compared, and some

contained ambiguous wording that was not optimal for extracting accurate measures of people's perceptions of moral decline. Moreover, all the items asked participants about the presence or absence of moral decline rather than asking them to rate the level of morality of people in both the present and the past. These limitations were addressed by studies 2a–c, but these studies had limitations of their own (for example, all participants were from the United States). And although studies 5a–b demonstrated the viability of the BEAM mechanism, they do not tell us whether it was the cause of the illusion of moral decline that our other studies documented.

With that said, the illusion of moral decline seems to be a robust phenomenon that may have troubling consequences. For example, in 2015, 76% of US Americans agreed that "addressing the moral breakdown of the country" should be a high priority for their government[27]. The United States faces many well-documented problems, from climate change and terrorism to racial injustice and economic inequality—and yet, most US Americans believe their government should devote scarce resources to reversing an imaginary trend. The belief that everyday morality is on the wane may also affect people's interpersonal behaviour. For example, research shows that people are reluctant to seek the aid and comfort of those whom they do not know because they underestimate how willingly those people would provide it[4,28,29]. The illusion of moral decline may be one of the reasons people do not depend as much as they might on the kindness of strangers—an act that might well ameliorate the illusion itself. The illusion of moral decline may also leave people dangerously susceptible to manipulation by bad actors. Research shows that people are especially influenced by 'dynamic norms', which are perceived changes in customary ways of behaving[5]. If low morality is a cause for concern, then declining morality may be a veritable call to arms, and leaders who promise to halt that illusory slide—to "make America great again", as it were—may have outsized appeal. Our studies indicate that the perception of moral decline is pervasive, perdurable, unfounded and easily produced. Achieving a better understanding of this phenomenon would seem a timely task.

Methods

Study 1

In study 1, we conducted keyword-term searches of the Roper Center for Public Opinion iPoll Database, and manually searched the databases of the General Social Survey, Pew Research Center, Gallup, the American National Election Studies, the World Values Survey, the European Social Survey and the European Values Survey to locate survey items that asked participants if and how they thought other people's morality had changed over time. In our analyses, we included all surveys that (1) used a representative sample of US American participants, and (2) explicitly asked participants about their perceptions of changes in values, traits and behaviours that have traditionally been taken as indicators of morality by a wide range of US Americans (for example, kindness, honesty, respect). We excluded from our analyses items that asked participants about their perceptions of special topics whose moral relevance either changed considerably over time (for example, men holding doors for women) or differed substantially across members of the population (for example, attending church). We also excluded items that asked participants about the morality of special subpopulations (for example, 'Evangelicals' or 'the Wisconsin legislature') rather than about all US Americans or about people in general. Further information, including search terms and all survey items included in study 1, can be found in the Supplementary Information. We also sampled our database for survey items administered to participants who lived outside the United States. Because there were fewer such surveys, we did not exclude surveys with non-representative samples, as we did with our US sample.

Study 2a

All original data collection in this and subsequent studies followed all ethical regulations and was approved by the Institutional Review Board of Harvard University.

Participants

We recruited a nationally representative sample of US American adults using Prolific, an online sample provider. This sample was constructed to represent the US American adult population in terms of gender, race and age. Because we did not know the size of the effect we were studying, we sought to make our sample comparable in size to the samples in study 1 by

recruiting 1,000 participants. Nine-hundred and ninety-nine people (507 female, 487 male, 5 other, M_{age} = 45.74 years, 73% white, 13% Black, 7% Asian, 4% Hispanic, 1% American Indian or Alaska Native, 1% other, 2% 'more than one of the above') were paid US$0.75 each for their participation.

Procedure

Study 2a was conducted in 2020. After providing informed consent, participants confirmed their Prolific ID, per the site's usage policy. They then read the following instructions: "Thanks! In this study, we'll ask you how kind, honest, nice, and good people were at various points in time. If you're not sure, that's okay, just give your best guess". Participants then rated how "kind, honest, nice, and good" people are today, were 10 years ago and were 20 years ago, using seven-point Likert scales with endpoints labelled 'not very' and 'very'. As a consistency check, participants were then asked to recall whether they had given higher, equal or lower ratings to people today compared to people 20 years ago. Participants then answered some open-ended exploratory questions that asked them to explain the thinking behind their answers. Participants then answered some demographic questions (Supplementary Table 6). Embedded among these demographic questions was an 'attention check question' that instructed participants to select the option 'other' and to type the word 'sky'. Finally, participants were compensated and dismissed.

Exclusions

One hundred and eighty-one participants failed the attention check embedded in the demographics and were excluded from all analyses. Another 120 participants gave answers to the consistency check question that were inconsistent with their previous answers; they were also excluded. This left 698 participants in all analyses (372 female, 322 male, four other, M_{age} = 46.37, 74% white, 12% Black, 6% Asian, 4% Hispanic, 1% American Indian or Alaska Native, 2% more than one of the above). These exclusions do not meaningfully affect the results.

Analysis

To analyse the data, we fit a linear mixed effects model using the lme4 package in R[30], extracted *P* values using the lmerTest package[31] and calculated planned contrasts using the emmeans package[32], using a Holm–Bonferroni correction for multiple comparisons. The outcome was participants' ratings and the predictor was the year of those ratings (one factor with three levels: 2020, 2010 and 2000). The model included a fixed effect of the year of each rating and a random intercept for each participant. For this and all models, we checked model assumptions by plotting the outcome variable, residuals and fitted values. All tests we report are two-tailed.

Study 2b

Participants

We powered study 2b to detect an effect of $d = 0.30$ or larger, reasoning that this would be sufficient to detect effects similar to the effect we detected in Study 2a. Two-hundred and thirty-six people responded to an advertisement for a study on Amazon Mechanical Turk. To participate, respondents had to pass a three-item test that required them to know that (1) children in kindergarten are 3 or 4 years old, (2) a US American ZIP code is a series of five digits and (3) eating turkey is not associated with Halloween. Thirty-six respondents answered at least one of these three questions incorrectly and were not allowed to participate. The remaining 200 respondents (81 female, 119 male, $M_{age} = 35.81$ years, 72% white, 12% Black, 9% Hispanic, 6% Asian, 3% more than one of the above) were allowed to participate in the study in exchange for US$0.75.

Procedure

After providing informed consent, participants followed study 2a's procedure except they were asked about different years. Specifically, participants were first asked, "How kind, honest, nice, and good are people today?" and were then asked the same question for "two years ago", "four

years ago", "six years ago", "eight years ago" and "ten years ago", in that order. All questions were answered using a seven-point Likert scale with endpoints labelled 'not very' and 'very'. As a consistency check, participants then answered the following question: "When it comes to being kind, honest, nice, and good—are people more so today compared to ten years ago, less so today compared to ten years ago, or the same?" Participants were then asked to explain their answer in an open-ended question. Finally, participants were asked some demographic questions, as well as an attention check question that required them to select the option 'other' and to type the word 'day'. Participants were compensated and dismissed.

Exclusions

Fifteen participants failed the attention check, and a further 37 participants failed the consistency check by giving an answer that was inconsistent with their scale ratings. The data from these participants were excluded from all analyses, leaving 148 participants (59 female, 89 male, $M_{age} = 36.59$ years, 75% white, 9% Black, 7% Hispanic, 5% Asian, 1% Hawaiian or Pacific Islander, 3% more the one of the above). These exclusions only meaningfully affect the results in one case, namely, that when all participants are included, the difference between 2020 and 2016 is not significant.

Analysis

We fit the same model we fit in study 2a except that in this case the factor in the model had six levels (2020, 2018, 2016, 2014, 2012 and 2010).

Study 2c

Participants

We sought to recruit a sample of people who varied widely in terms of age. As such, we created a survey with a quota of 50 participants in each of the following age groups: 18–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64 and 65–69 years. This sample size gave us sufficient power to

detect the effects we had detected in studies 2a and 2b. Respondents selected their age group on accessing the study, and once the quota for a group was reached, further respondents from that group were not allowed to participate. Respondents younger than 18 or older than 69 were not allowed to participate.

Respondents responded to an advertisement for a study on Amazon Mechanical Turk. Respondents who accessed the survey before the quota for their age group was reached were asked to complete a three-item test of English proficiency and knowledge of US American culture. Specifically, they were required to demonstrate that they knew that (1) bell bottoms are not a type of footwear, (2) an RSVP is a required response to a wedding invitation and (3) a sign reading ‘out of order’ is best paired with an elevator. Three hundred and one respondents answered one or more of these questions incorrectly and were not allowed to participate. The remaining 484 respondents (225 female, 257 male, two other, $M_{age} = 41.27$ years, 72% white, 15% Black, 7% Asian, 4% Hispanic, 1% American Indian or Alaska Native, 2% more than one of the above) were allowed to participate in the study in exchange for US$0.75.

Procedure

Study 2c was conducted in 2020. Participants responded to an advertisement for a study on Amazon Mechanical Turk. After providing informed consent, participants reported how “kind, honest, nice and good” people are today. They then reported how “kind, honest, nice and good” people were when they (the participants) were about 20 years old, and at about the time they (the participants) were born. This was done by adjusting the wording of the subsequent questions on the basis of the participant’s age. For example, if the participant was between 30 and 34 years old, they were asked “How kind, honest, nice, and good were people about ten years ago?” and then “How kind, honest, nice, and good were people about 30 years ago?” If participants were under 25 years, they answered only the questions for today and when they were born. All questions were answered using a seven-point Likert scale with endpoints labelled ‘not very’ and ‘very’. As in previous studies, participants were then given a consistency check that required them to remember whether they had rated people today as more, equally or less

moral compared to people in the year they were born. Participants then answered some further exploratory and demographic questions. Embedded among them was an attention check that required participants to select the option ‘other’ and type the word ‘apple’. Finally, participants were compensated and dismissed.

Exclusions

Twenty-eight participants failed the attention check and their data were excluded from all analyses. Seventy-three more participants reported an age at the end of the study that was inconsistent with the age group they selected at the beginning of the study and the data from these participants were also excluded from all analyses. An extra 64 participants failed the consistency check and data from these participants were also excluded from all analyses. The data from the remaining 347 participants (174 female, 172 male, one other, $M_{age} = 42.57$ years, 78% white, 9% Black, 7% Asian, 4% Hispanic, 2% ‘more than one of the above’) were included in all analyses. These exclusions do not meaningfully change the results.

Analysis

We fit the same model we fit in study 2b except that in this case the factor in the model had three levels (today, the year the participant turned 20, the year the participant was born).

Study 3

Participants

Respondents responded to an advertisement for a study on Amazon Mechanical Turk. As in study 2c, we sought to recruit a sample of people who varied widely in terms of age and that was large enough to provide sufficient power to detect the effects we had detected in studies 2a and 2b. We created a survey with quota of 150 for each of three age groups: 20–34, 35–49 and 50–64. Anyone younger than 20 or older than 64 was not allowed to participate. Respondents were asked to complete the same test of English

language and US American culture as in study 2c. Four hundred and forty-four respondents (202 female, 242 male, M_{age} = 40.42 years, 77% white, 9% Black, 7% Asian, 5% Hispanic, 1% ‘more than one of the above’) provided informed consent and became participants in the study in exchange for US$0.75.

Procedure

Study 3 was conducted in 2020. After providing informed consent, participants reported how “kind, honest, nice, and good” people are in the present (2020) and also “about 15 years ago” (about 2005) on seven-point Likert scales with endpoints labelled ‘not very’ and ‘very’ and then completed a consistency check that asked them to recall the answers they had just given. The difference between these two ratings was used as a measure of participants’ perception of moral decline between 2005 and 2020. Participants then answered the following questions using the same seven-point Likert scales: “How kind, honest, nice, and good are people who are currently between the ages of 35 and 95?”; “How kind, honest, nice, and good are people who are currently between the ages of 20 and 35?”; “Thinking again of people who are currently between the ages of 35 and 95, how kind, honest, nice, and good were they about 15 years ago?” and “About 15 years ago, how kind, honest, nice, and good were people who were then between the ages of 80 and 95?” Participants then answered some demographics questions, among which was embedded an ‘attention check question’ that instructed participants to select the option ‘other’ and to type the word ‘cloud’. Finally, participants were compensated and dismissed.

Exclusions

Forty-eight participants failed the attention check, and a further 15 participants reported an age at the end of the study that was inconsistent with the age group they reported at the beginning of the study. An extra 77 participants failed the consistency check. The data from all of these participants were excluded from all analyses, leaving 319 participants (154 female, 165 male, M_{age} = 41.02, 77% white, 8% Black, 8% Asian, 5%

Hispanic, 1% more than one of the above). These exclusions do not meaningfully affect the results.

Calculating personal change and interpersonal replacement

We created a personal change score by subtracting ratings of 20–80-year olds about 15 years ago (in 2005) from ratings of 35–95-year olds in the present (2020). We created an interpersonal replacement score by subtracting ratings of 80–95-year olds about 15 years ago (in 2005) from ratings of 20–35-year olds in the present (2020). The descriptive statistics for people in general and each of the subgroups about which participants were asked are shown in Extended Data Fig. 1.

Analysis

Using a standard linear model, we entered participants' personal change and cohort replacement scores as predictors, and the outcome was participants' overall perception of moral decline between 2005 and 2020.

Study 4

In study 4, we conducted keyword-term searches of the Roper Center for Public Opinion iPoll Database (using search terms shown in the Supplementary Information), and manually searched the databases of the General Social Survey, Pew Research Center, Gallup, the American National Election Studies, the World Values Survey, the European Social Survey and the European Values Survey to locate survey items that asked participants questions about their own and other people's morality. As in study 1, questions were considered relevant to morality if they asked about values, attitudes, traits and behaviours that we thought would be considered relevant to kindness, honesty, niceness and goodness by a wide range of US Americans. We included US samples only if they were nationally representative, but also collected non-representative samples if they were collected outside the United States to maximize non-US representation. The latter were analysed separately. To be included, each survey had to be administered at least twice, and the most recent administration could not be

earlier than 2010. Further information, including search terms and all survey items included in study 4, can be found in the Supplementary Information.

Analysis

We fit a linear model for each survey. The year of each survey was always entered as a predictor and the outcome was always the average perception of current morality. We used R^2 values as a measure of effect size. We fit Bayesian models using the Rstanarm package in R[33] and extracted the percentage of the 89% HDI that was contained in the ROPE, which was by default defined as ±0.1 standard deviations. We used the package's default Markov Chain Monte Carlo and prior settings ($M = 0$, scale of 2.5).

Study 5a

Participants

As in study 2c, we sought to recruit a sample of people who varied widely in terms of age and that was large enough to provide sufficient power to detect the effects we had detected in previous studies. We created a survey with a quota of 50 participants in each of three age groups: 20–34, 35–49 and 50–64 years. Anyone who was either younger than 20 years or older than 64 years was not allowed to participate.

One thousand and twenty-one people responded to an advertisement for a study on Amazon Mechanical Turk. They completed the same test of English language and US American culture as in study 2c. Five hundred and twenty-one respondents answered at least one of the questions incorrectly and were not allowed to participate. The remaining 500 respondents (204 female, 293 male, three other, $M_{age} = 37.74$ years, 65% white, 24% Black, 7% Asian, 2% Hispanic, 1% American Indian or Alaska Native, 1% more than one of the above) provided informed consent and became participants in the study in exchange for US$0.75.

Procedure

Study 5a was conducted in 2021. After providing informed consent, participants completed the same procedure as was used in study 2c, with two more questions. Specifically, participants rated how "kind, honest, nice, and good" people in general were 20 years before the participant was born and also 40 years before the participant was born. These years were adjusted on the basis of the age of the participant.

Exclusions

One hundred and seventy-nine participants failed the first attention check, and another 21 failed the second attention check. Another 15 participants reported an age at the end of the study that was inconsistent with the birth year they reported at the beginning. The data from all these participants were excluded from all analyses. The remaining 283 participants (139 female, 143 male, one other, $M_{age} = 38.77$ years, 78% white, 11% Black, 8% Asian, 2% Hispanic, 1% more than one of the above) were included in all analyses. These exclusions affect the results in a few cases. Specifically, when excluded participants are included, the overall perception of moral decline and personal change for people in general are not significant. All other effects remain significant.

Analysis

We fit the same model we fit in study 2c except that the factor in the model had five levels (2020, the year the participant turned 20, the year the participant was born, 20 years before the participant was born and 40 years before the participant was born).

Study 5b

Participants

Because this study was a replication and extension of study 2c, we sought to collect a similar sample size to have the power to detect similar effects, and we used the same age quotas as in Study 2c. One thousand eighty-two people responded to an advertisement for a study on Amazon Mechanical

Turk. Twenty-one of these opened the study but did not complete it. Five hundred and sixty people responded after the quota for their age group had been reached and were not allowed to participate in the study. Respondents who responded before the quota for their age group was reached completed the same three-item test of US American culture and English language used in study 2c. Twenty-three respondents answered one or more of these questions incorrectly and were not allowed to participate in the study. The remaining 499 respondents (225 female, 241 male, three other, $M_{age} = 43.96$ years, 78% white, 10% Asian, 5% Black, 4% Hispanic, 3% more than one of the above) were allowed to participate in the study in exchange for US$0.75.

Procedure

Study 5b was conducted in 2021. After providing informed consent, participants completed the same procedure used in study 2c. They further rated people's morality 20 and 40 years before the year that they were born.

Exclusions

Forty-four participants failed the attention check and their data were excluded from all analyses. Seven more participants reported an age at the end of the study that was inconsistent with the age group they selected at the beginning of the study and their data were also excluded from all analyses. Sixty-one more participants failed the consistency check and their data were also excluded from all analyses. The data from the remaining 387 participants (206 female, 178 male, three other, $M_{age} = 44.04$ years, 79% white, 11% Asian, 4% Black, 3% Hispanic, 2% more than one of the above) were included in all analyses. These exclusions affect the results in one case: when excluded participants are included, participants perceived moral improvement from 40 years before birth to 20 years before birth. All other effects remain the same.

Analysis

We fit the same model we fit in study 2c except that in this case the factor in the model had five levels (the year 2020, the year the participant turned 20 years old, the year the participant was born, 20 years before the participant was born and 40 years before the participant was born).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All materials and original data are available at https://osf.io/t83zy/ (https://doi.org/10.17605/OSF.IO/T83ZY). The data analysed in studies 1 and 4 are the property of the polling organizations that produced them and cannot be posted. Instructions for accessing these data are also available at https://osf.io/t83zy/.

Code availability

The code necessary to reproduce all analyses is available at https://osf.io/t83zy/, except for studies 1 and 4, which requires access to proprietary data.

References

1. Eibach, R. P. & Libby, L. K. In *Social and Psychological Bases of Ideology and System Justification* (ed. Jost, J. T. et al.) 402–423 (Oxford Univ. Press, 2009).

2. Herman, A. *The Idea of Decline in Western History* (Free Press, 1997).

3. Thaler, R. H. & Sunstein, C. R. *Nudge: Improving Decisions about Health, Wealth and Happiness* (Penguin Books, 2009).

4. Epley, N., Kardas, M., Zhao, X., Atir, S. & Schroeder, J. Undersociality: miscalibrated social cognition can inhibit social connection. *Trends Cogn. Sci.* **26**, 406–418 (2022).

5. Sparkman, G. & Walton, G. M. Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychol. Sci.* **28**, 1663–1674 (2017).

6. Conway, R. S. & Walters, C. F. *Titus Livius (Livy) Ab urbe condita (History of Rome)* (Oxford Univ. Press, 1914).

7. Murphy, A. R. Augustine and the rhetoric of Roman decline. *Hist. Polit. Thought* **26**, 586–606 (2005).

8. Ellemers, N., van der Toorn, J., Paunov, Y. & van Leeuwen, T. The psychology of morality: a review and analysis of empirical studies published from 1940 through 2017. *Personal. Soc. Psychol. Rev.* **23**, 332–366 (2019).

9. Marsh, A. A. et al. Neural and cognitive characteristics of extraordinary altruists. *Proc. Natl Acad. Sci. USA* **111**, 15036–15041 (2014).

10. Wygant, D. B., Pardini, D. A., Marsh, A. A. & Patrick, C. J. in *Handbook of Psychopathy* 2nd edn (ed. Patrick, C. J.) 755–778 (The Guilford Press, 2018).

11. *Summer 2002 Survey Data* (Pew Research Center, 2002).

12. *Spirit and Power – A 10-Country Survey of Pentecostals* (Pew Research Center, 2006).

13. Protzko, J. & Schooler, J. W. Who denigrates today's youth?: the role of age, implicit theories, and sharing the same negative trait. *Front. Psychol.* **13**, 723515 (2022).

14. Protzko, J. & Schooler, J. W. Kids these days: why the youth of today seem lacking. *Sci. Adv.* **5**, eaav5916 (2022).

15. Pinker, S. *The Better Angels of our Nature: Why Violence has Declined* (Viking, 2011).

16. Pinker, S. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress* (Viking, 2018).

17. Yuan, M. et al. Did cooperation among strangers decline in the United States? A cross-temporal meta-analysis of social dilemmas (1956–2017). *Psychol. Bull.* **148**, 129–157 (2022).

18. Baumeister, R. F., Bratslavsky, E., Finkenauer, C. & Vohs, K. D. Bad is stronger than good. *Rev. Gen. Psychol.* **5**, 323–370 (2001).

19. Rozin, P. & Royzman, E. B. Negativity bias, negativity dominance, and contagion. *Personal. Soc. Psychol. Rev.* **5**, 296–320 (2001).

20. Pratto, F. & John, O. P. Automatic vigilance: the attention-grabbing power of negative social information. *J. Pers. Soc. Psychol.* **61**, 380–391 (1991).

21. Soroka, S. & McAdams, S. News, politics, and negativity. *Polit. Commun.* **32**, 1–22 (2015).

22. Breslin, C. W. & Safer, M. A. Effects of event valence on long-term memory for two baseball championship games. *Psychol. Sci.* **22**, 1408–1412 (2011).

23. D'Argembeau, A. & Van der Linden, M. Remembering pride and shame: self-enhancement and the phenomenology of autobiographical memory. *Memory* **16**, 538–547 (2008).

24. Mitchell, T. R., Thompson, L., Peterson, E. & Cronk, R. Temporal adjustments in the evaluation of events: the 'rosy view'. *J. Exp. Soc. Psychol.* **33**, 421–448 (1997).

25. Skowronski, J. J., Walker, W. R., Henderson, D. X. & Bond, G. D. in *Advances in Experimental Social Psychology* Vol. 49 (eds. Olson, J. M. & Zanna, M. P.) 163–218 (Elsevier, 2014).

26. Lüdecke, D., Ben-Shachar, M., Patil, I. & Makowski, D. Extracting, computing and exploring the parameters of statistical models using R. *J. Open Source Softw.* **553**, 2445 (2020).

27. Pew Research Center for the People & the Press Research Center Poll: January 2015 Political Survey (Version 2). *Roper Center for Public Opinion Research* https://doi.org/10.25940/ROPER-31096284 (2015).

28. Zhao, X. & Epley, N. Surprisingly happy to have helped: underestimating prosociality creates a misplaced barrier to asking for help. *Psychol. Sci.* https://doi.org/10.1177/09567976221097615 (2022).

29. Dungan, J. A., Munguia Gomez, D. M. & Epley, N. Too reluctant to reach out: receiving social support is more positive than expressers expect. *Psychol. Sci.* https://doi.org/10.1177/09567976221082942 (2022).

30. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

31. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* https://doi.org/10.18637/jss.v082.i13 (2017).

32. Lenth, R. emmeans: estimated marginal means, aka least-squares means. R package version 1.4.7 (R Foundation for Statistical Computing, 2020).

33. Goodrich, B., Gabry, J., Ali, I. & Brilleman, S. rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1 (R Foundation for Statistical Computing, 2020).

Acknowledgements

We thank P. Menon for her mathematical assistance, M. Vollberg for his coding assistance, D. Blumenthal for his assistance with data collection and F. Cushman, N. Epley, S. Pinker, R. Baumeister, A. Wood Brooks and T. Wilson for comments on the manuscript.

Author information

Authors and Affiliations

1. Columbia University, New York, NY, USA

   Adam M. Mastroianni

2. Harvard University, Cambridge, MA, USA

   Daniel T. Gilbert

Contributions

A.M.M. collected and analysed the data. A.M.M. and D.T.G. developed the study concepts, developed the study designs, drafted the manuscript and approved the manuscript.

Corresponding author

Correspondence to Adam M. Mastroianni.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Roy Baumeister and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Descriptive Statistics in Study 3.

Error bars represent 95% confidence intervals. Transparent points represent individual observations jittered for legibility.

Supplementary information

Supplementary Information

Reporting Summary

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Mastroianni, A.M., Gilbert, D.T. The illusion of moral decline. *Nature* **618**, 782–789 (2023). https://doi.org/10.1038/s41586-023-06137-x

- Received: 11 July 2022
- Accepted: 26 April 2023
- Published: 07 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06137-x

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

[Morality is declining, right? Scientists say that idea is an illusion](https://www.nature.com)

- Mariana Lenharo

Nature News 07 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06137-x

- Article
- Open Access
- Published: 14 June 2023

Psychedelics reopen the social reward learning critical period

- Romain Nardou,
- Edward Sawyer,
- Young Jun Song,
- Makenzie Wilkinson,
- Yasmin Padovan-Hernandez,
- Júnia Lara de Deus,
- Noelle Wright,
- Carine Lama,
- Sehr Faltin,
- Loyal A. Goff,
- Genevieve L. Stein-O'Brien &
- …
- Gül Dölen

Show authors

Nature **volume 618**, pages 790–798 (2023)

- 87k Accesses

- 1 Citations

- 779 Altmetric

- Metrics details

Abstract

Psychedelics are a broad class of drugs defined by their ability to induce an altered state of consciousness[1,2]. These drugs have been used for millennia in both spiritual and medicinal contexts, and a number of recent clinical successes have spurred a renewed interest in developing psychedelic therapies[3,4,5,6,7,8,9]. Nevertheless, a unifying mechanism that can account for these shared phenomenological and therapeutic properties remains unknown. Here we demonstrate in mice that the ability to reopen the social reward learning critical period is a shared property across psychedelic drugs. Notably, the time course of critical period reopening is proportional to the duration of acute subjective effects reported in humans. Furthermore, the ability to reinstate social reward learning in adulthood is paralleled by metaplastic restoration of oxytocin-mediated long-term depression in the nucleus accumbens. Finally, identification of differentially expressed genes in the 'open state' versus the 'closed state' provides evidence that reorganization of the extracellular matrix is a common downstream mechanism underlying psychedelic drug-mediated critical period reopening. Together these results have important implications for the implementation of psychedelics in clinical practice, as well as the design of novel compounds for the treatment of neuropsychiatric disease.

Main

Classically, psychedelics have been defined to include drugs such as lysergic acid diethylamide (LSD), mescaline, phenylcyclohexyl piperidine (PCP), ibogaine, 3,4-methylenedioxymethamphetamine (MDMA), psylocibin and ketamine, because each of these compounds produces alterations to sensory, self, time and space perception that are "so alien to everyday experience that they shed new light on the workings of these everyday mental functions"[1]. Although more recent attempts have been made to subcategorize psychedelics[10] on the basis of the subjective character of the altered state that they induce (for example, hallucinogenic, empathogenic, oneirogenic or dissociative), their chemical structure (for example, tryptamines, phenethylamines or arylcyclohexamines), or their principal binding target (for example, serotonin receptor 2A ($5\text{-}HT_{2A}R$), monoamine transporter, κ-

opioid receptor (KOR) or *N*-methyl-d-aspartate receptor (NMDAR)), the importance of these categories for therapeutic applications remains unclear, since psychedelics that span the diversity of classification systems have shown remarkable promise for the treatment of addiction[4,5], post-traumatic stress disorder[6,7] (PTSD) and depression[3,8,9]. Thus, identification of a common neurobiological mechanism that can account for the shared therapeutic effects of psychedelics is an obvious priority for translational neuroscience.

During specific periods of brain development, the nervous system exhibits heightened sensitivity to ethologically relevant stimuli, as well as increased malleability for synaptic, circuit and behavioural modifications. These mechanistically constrained windows of time are called critical periods and neuroscientists have long sought methods to reopen them for therapeutic benefit. Recently, we have discovered a novel critical period for social reward learning and shown that the empathogenic psychedelic MDMA is able to reopen this critical period[11]. This mechanism shares a number of features with the therapeutic effects of MDMA-assisted psychotherapy for the treatment of PTSD, including rapid onset, durability and context dependence[6,7]. At the same time, cocaine does not reopen the social reward learning critical period[11], and since cocaine does not share the psychedelics' therapeutic profile[12], these results lend further support for the view that the reinstatement of social reward learning in adulthood underlies the therapeutic efficacy of MDMA.

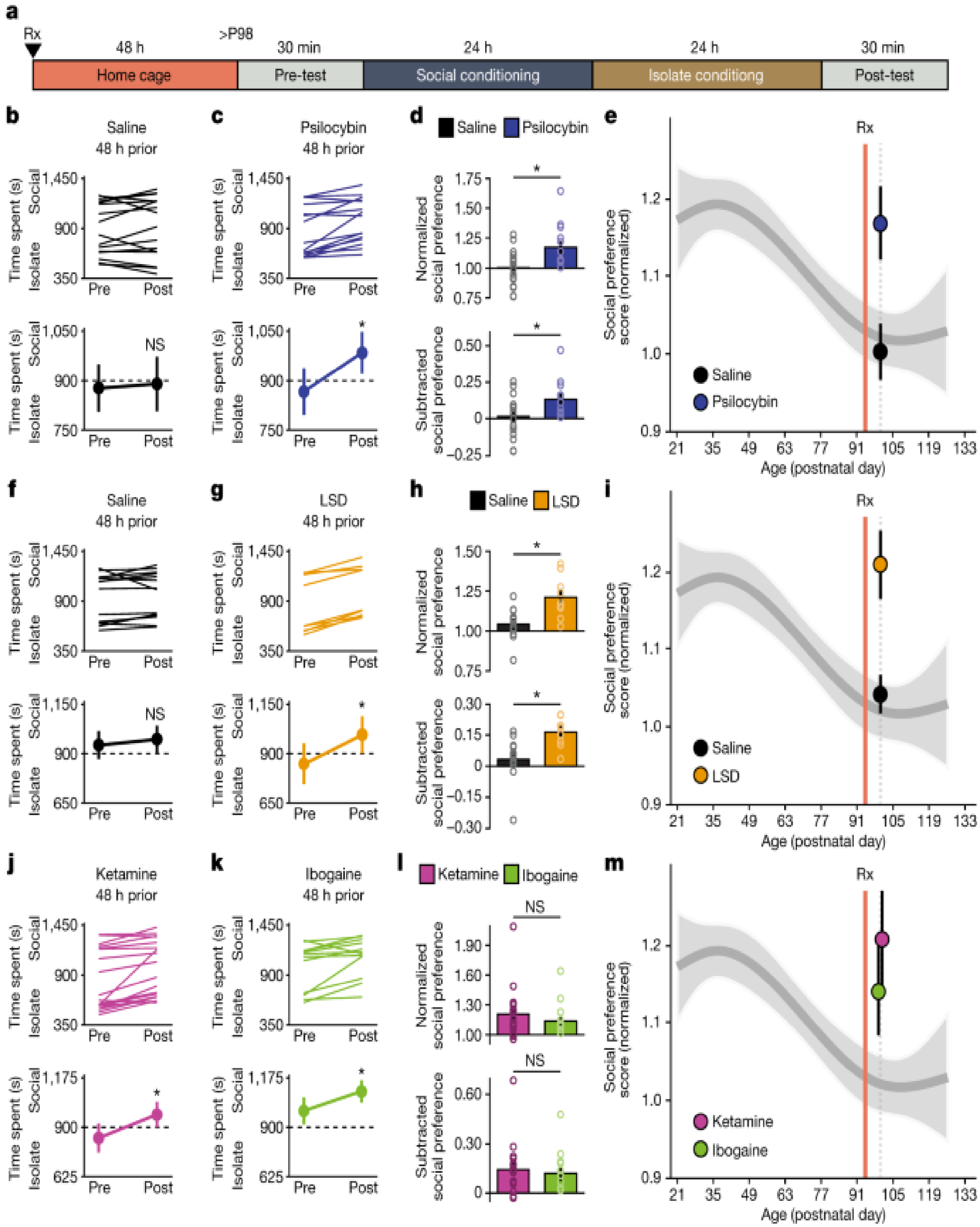
Whether the ability of MDMA to reopen the critical period for social reward learning generalizes across psychedelics remains an open question. MDMA is classified as an 'empathogen' because its acute subjective effects are distinctly prosocial in quality[13]. The fact that this quality is not shared by hallucinogenic psychedelics such as psilocybin and LSD[14], dissociative psychedelics such as ketamine[15], or oneirogenic psychedelics such as ibogaine[16] challenges the idea that these drugs could reopen the social reward learning critical period. However, the psychotropic effects of MDMA include an altered state of consciousness shared by all psychedelics[1,2], and if it is this characteristic rather than its prosocial properties that embodies the subjective experience of reopening critical

periods, then the ability to reinstate social reward learning in adulthood might generalize across psychedelics.

Critical period reopening is a shared property

To test whether the ability of MDMA to reopen the social reward learning critical period generalizes across psychedelics, we began by examining the effect of psilocybin pretreatment on the magnitude of social reward learning in adulthood using the social reward conditioned place preference (sCPP) assay (Extended Data Fig. 1). We administered a single intraperitoneal (i.p.) dose of psilocybin[17] (0.3 mg kg^{-1}) to adult male mice (at postnatal day 96 (P96)) and 48 h later (at P98), we assessed the magnitude of sCPP (Fig. 1a). Mice pretreated with psilocybin, but not saline, exhibited a significant sCPP at P98 (Fig. 1b–d). To formally designate 'open' and 'closed' states of this critical period, we next generated a natural spline regression model to previously published data[11] with knots at P35 and P98 ($P = 1.003 \times 10^{-6}$; root mean square error (r.m.s.e.) = 0.19; $R^2 = 0.11$), as shown in Extended Data Fig. 2. When compared with this derived curve, the magnitude of sCPP in saline-treated mice did not deviate significantly from the closed state ($P = 0.72$), whereas the fit derived from psilocybin-treated mice demonstrated a significant mean shift ($P = 1.12 \times 10^{-6}$) in range of the open state (Fig. 1e). Similarly, pretreatment with LSD[17] (i.p. 1 μg kg^{-1}) but not saline, also reopened the critical period for social reward learning (saline $P = 0.90$, LSD $P = 1.76 \times 10^{-9}$) (Fig. 1f–i). Next, we examined the effects of ketamine[18] (i.p. 3 mg kg^{-1}) and ibogaine[19] (i.p. 40 mg kg^{-1}). Mice pretreated with either drug also exhibited sCPP in adulthood ($P = 8.78 \times 10^{-4}$ and $P = 3.17 \times 10^{-5}$, respectively) (Fig. 1k–m). As with MDMA[11], these effects were dose-dependent (Extended Data Fig. 3). In juveniles, MDMA[11] (i.p. 10 mg kg^{-1}) pretreatment did not lead to a further increase the magnitude of social reward learning (Extended Data Fig. 3). In contrast to its effects on social reward learning behaviour, pretreatment with psychedelics had no effect on the magnitude of two addiction-like behaviours: cocaine reward learning and amphetamine-induced locomotor sensitization (Extended Data Fig. 4). Together, these studies demonstrate that as with empathogenic psychedelics[11], hallucinogenic, oneirogenic and dissociative psychedelics are able to reopen the critical period for social reward learning.

Fig. 1: Psychedelics reopen the social reward learning critical period.



a, Experimental time course of i.p. pretreatment (Rx) in sCPP. **b**,**c**,**f**,**g**,**j**,**k**, Individual (top) and average (bottom) responses of P98 mice indicate that mice pretreated with psilocybin (0.3 mg kg^{-1}) (**c**; $n = 15$, $t_{(14)} = -3.741$, $P =$

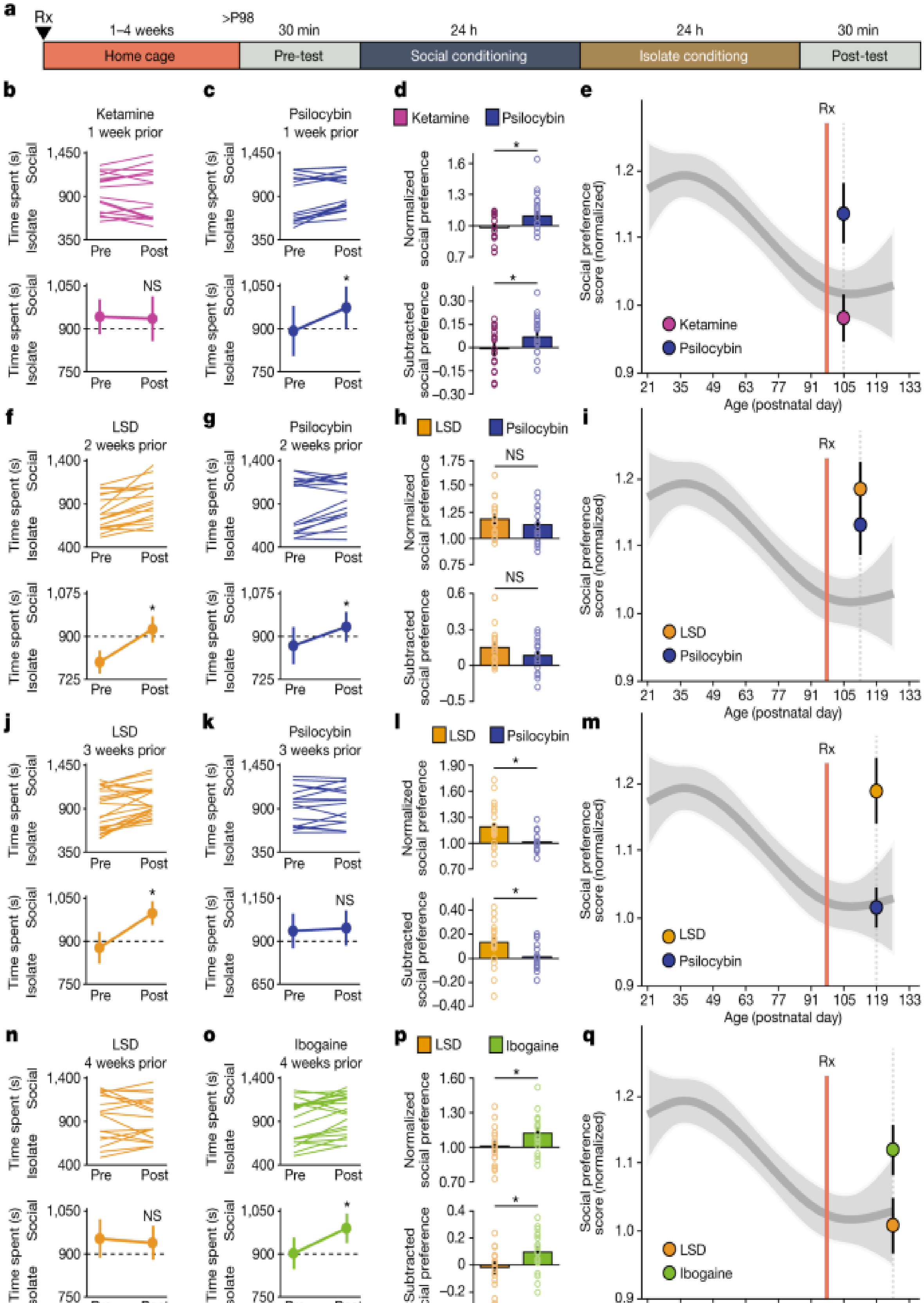
0.002), LSD (1 μg kg^{-1}) (**g**; $n = 9$, $t_{(8)} = -7.095$, $P < 0.001$), ketamine (3 mg kg^{-1}) (**j**; $n = 18$, $t_{(17)} = -3.826$, $P < 0.002$), and ibogaine (40 mg kg^{-1}) (**k**; $n = 12$, $t_{(11)} = -2.690$, $P = 0.02$) but not saline (**b**; $n = 17$ mice, $t_{(16)} = -0.441$, $P = 0.665$. **f**; $n = 14$ mice, $t_{(13)} = -1.215$, $P = 0.25$) develop a preference for the social bedding cue. Two-tailed paired *t*-test. **d**,**h**,**l**, Comparisons reveal a significant increase in normalized (top) and subtracted (bottom) social preference for pretreatment with psilocybin versus saline (**d**; normalized, $t_{(30)} = -2.800$, $P = 0.009$; subtracted, $t_{(30)} = -2.401$, $P = 0.023$), and with LSD versus saline (**h**; normalized, $t_{(21)} = -3.558$, $P = 0.002$; subtracted, $t_{(21)} = -3.344$, $P = 0.003$), but no difference between pretreatment with ketamine and ibogaine (**l**; normalized, $t_{(28)} = 0.749$, $P = 0.460$; subtracted, $t_{(28)} = 0.409$, $P = 0.686$). Two-tailed unpaired *t*-test, with Welch's correction to account for unequal variance in **l** subtracted. *$P < 0.05$; NS, not significant ($P > 0.05$). **e**,**i**,**m**, Normalized social preference in mice pretreated with psilocybin versus saline (**e**), LSD versus saline (**i**) and ibogaine versus ketamine (**m**), plotted against a natural spline regression model of the developmental time course of normalized social preference scores. Comparison with the natural spline model revealed that the magnitude of sCPP in saline-treated mice did not deviate significantly from the closed state (**b**; $P = 0.72$) (**f**; $P = 0.90$), whereas mice pretreated with psilocybin ($P = 1.12 \times 10^{-6}$), LSD ($P = 1.76 \times 10^{-9}$), ketamine ($P = 8.78 \times 10^{-4}$) or ibogaine ($P = 3.17 \times 10^{-5}$) demonstrated a significant mean shift in range of the open state. Comparisons with the natural spline model were considered not significant ($P > 0.1$). Rx indicates drug treatment. Data are as mean ± s.e.m. n refers to the number of biologically independent mice.

Duration of the psychedelic open state

The duration of acute subjective effects and the durability of the therapeutic response vary considerably across psychedelics. For example, in humans, the acute subjective effects of ketamine[15] last 30–120 min, whereas its antidepressant effects[9] last for 1 week. By contrast, the subjective effects of psilocybin and MDMA[20,21] last for 3–6 h, whereas the acute effects of LSD and ibogaine persist for 8–10 h and 36–72 h, respectively[16,22]; these long-lasting subjective effects correspond to highly durable therapeutic effects

that last months to years[4,5,7]. Previously, we showed that MDMA-induced critical period reopening lasts for two weeks, but returns to the closed state by four weeks[11]. Here, to further probe the time course of the critical period open state induced by psychedelics, we examined the duration of critical period reopening following treatment with ketamine, psilocybin, LSD and ibogaine (Fig. 2a). One week following psychedelic treatment, psilocybin-treated mice, but not those treated with ketamine, exhibited significant social reward learning (Fig. 2b–e). Two weeks following psychedelic treatment, the social reward learning critical period remained open for both psilocybin- and LSD-treated mice (Fig. 2f–i). At three weeks, LSD-treated mice, but not those treated with psilocybin, exhibited significant social reward learning (Fig. 2j–m), whereas at four weeks, the social reward learning critical period remained open for mice treated with ibogaine but not those treated with LSD (Fig. 2n–q). For each psychedelic, we examined at least three time points; increasing the LSD dose to 50 $\mu g\,kg^{-1}$ did not extend the duration of the open state (Extended Data Fig. 5). As shown in Fig. 3, the progressively longer-lasting open states induced by ketamine (Figs. 1f–i and 2b–e and Extended Data Fig. 5), followed by psilocybin (Fig. 2b–i), MDMA[11] (Extended Data Fig. 5), LSD (Fig. 2j–q) and ibogaine (Fig. 2n–q and Extended Data Fig. 5) are proportional to the duration of the acute subjective effects of these drugs in humans[15,16,20,21,22]. These results provide a mechanistic explanation for the importance of the post-treatment integration period for clinical implementation of psychedelics, and inform the design of novel compounds for clinical applications.

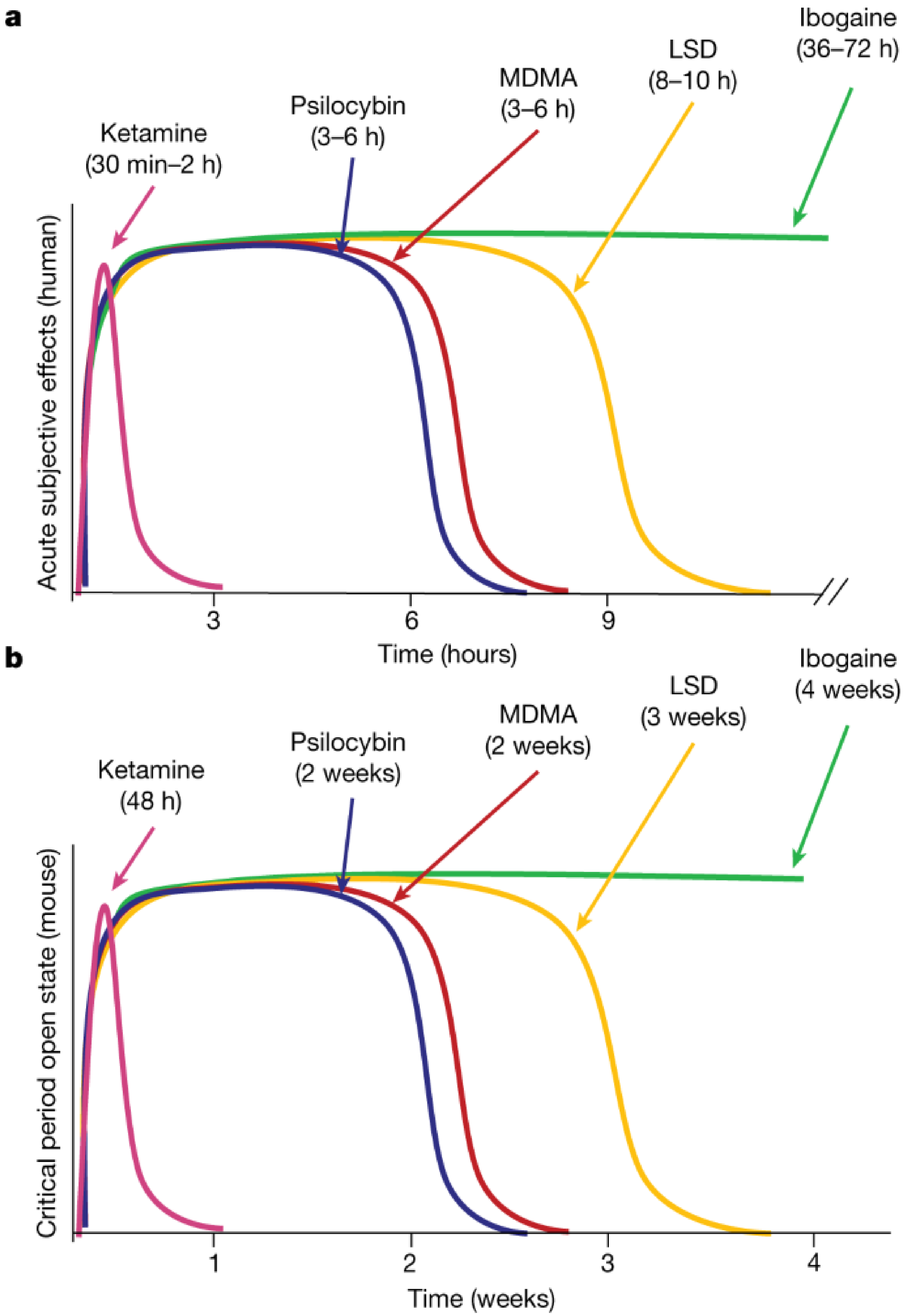
Fig. 2: The duration of the open state induced by psychedelics is variable.


a
Rx
1–4 weeks
>P98
30 min
24 h
24 h
30 min
Home cage
Pre-test
Social conditioning
Isolate conditiong
Post-test
b
Ketamine
1 week prior
c
Psilocybin
1 week prior
d
Ketamine
Psilocybin
e
Rx
Ketamine
Psilocybin
f
LSD
2 weeks prior
g
Psilocybin
2 weeks prior
h
LSD
Psilocybin
i
LSD
Psilocybin
j
LSD
3 weeks prior
k
Psilocybin
3 weeks prior
l
LSD
Psilocybin
m
LSD
Psilocybin
n
LSD
4 weeks prior
o
Ibogaine
4 weeks prior
p
LSD
Ibogaine
q
LSD
Ibogaine
Time spent (s)
Social
Isolate
Pre
Post
NS
*
Normalized social preference
Subtracted social preference
Social preference score (normalized)
Age (postnatal day)

a, Experimental time course of i.p. pretreatment in the sCPP assay. **b**–**q**, sCPP in adult mice 1 week after i.p. pretreatment with ketamine (3 mg kg^{-1}) or psilocybin (0.3 mg kg^{-1}) (**b**–**e**), 2 weeks after pretreatment with LSD (1 μg kg^{-1}) or psilocybin (0.3 mg kg^{-1}) (**f**–**i**), 3 weeks after pretreatment with LSD (1 μg kg^{-1}) or psilocybin (0.3 mg kg^{-1}) (**j**–**m**) or 4 weeks after pretreatment with LSD (1 μg kg^{-1}) or ibogaine (40 mg kg^{-1}) (**n**–**q**). **b**,**c**,**f**,**g**,**j**,**k**,**n**,**o**, Individual (top) and average (bottom) responses indicate the reinstatement of sCPP is absent one week after ketamine treatment (**b**, $n = 16$ mice, $t_{(15)} = 0.204$, $P = 0.841$), lasts two weeks for psilocybin (**c**, 1 week: $n = 17$ mice, $t_{(16)} = -2.959$, $P = 0.009$; **g**, 2 weeks: $n = 22$ mice, $t_{(21)} = -3.542$, $P = 0.002$; **k**, 3 weeks: $n = 16$ mice, $t_{(15)} = -0.405$, $P = 0.691$), lasts 3 weeks for LSD (**f**, 2 weeks: $n = 18$ mice, $t_{(17)} = -4.360$, $P < 0.001$; **j**, 3 weeks: $n = 23$ mice, $t_{(22)} = -3.671$, $P = 0.001$; **n**, 4 weeks: $n = 17$ mice, $t_{(16)} = 0.441$, $P = 0.665$), and lasts at least 4 weeks for ibogaine (**o**, $n = 20$ mice, $t_{(19)} = -3.004$, $P = 0.007$). Two-tailed paired *t*-test. **d**,**h**,**l**,**p**, Comparisons reveal a significant difference in sCPP between ketamine and psilocybin groups 1 week after pretreatment (**d**, normalized: $t_{(31)} = -2.700$, $P = 0.011$; subtracted: $t_{(31)} = -2.113$, $P = 0.043$), between LSD and psilocybin at 3 weeks (**l**, normalized: $t_{(34)} = 3.050$, $P = 0.004$; subtracted: $t_{(37)} = 2.471$, $P = 0.018$) but not at 2 weeks (**h**, normalized: $t_{(38)} = 0.390$, $P = 0.699$; subtracted: $t_{(38)} = 1.077$, $P = 0.288$), and LSD and ibogaine 4 weeks after pretreatment (**p**, normalized: $t_{(35)} = -2.045$, $P = 0.048$; subtracted: $t_{(35)} = -2.283$, $P = 0.029$). Two-tailed unpaired *t*-test, with Welch's correction to account for unequal variance in **l** subtracted. *$P < 0.05$; NS, not significant ($P > 0.05$). **e**,**i**,**m**,**q**, Normalized social preference one week after ketamine or psilocybin (**e**), two (**i**) and three (**m**) weeks after LSD and psilocybin, and four weeks after LSD and ibogaine (**q**) plotted against a natural spline model of the developmental time course of normalized social preference scores. The magnitude of sCPP did not deviate significantly from the closed state 1 week after ketamine (**e**, $P = 0.949$), three weeks after psilocybin (**i**, $P = 0.633$) and four weeks after LSD (**m**, $P = 0.705$), whereas the magnitude demonstrated a significant mean shift in range of the open state for both one (**e**, $P = 0.054$) and two weeks (**i**, $P = 0.0211$) after psilocybin, two (**i**, $P = 0.0121$) and three weeks (**m**, $P = 0.00745$) after LSD and four weeks after ibogaine (**q**, $P = 0.0758$). Comparisons to the natural spline model were considered not significant (P

> 0.1). Data are mean ± s.e.m. *n* refers to the number of biologically independent mice.

Fig. 3: The durations of acute subjective effects in humans are proportional to the durations of the critical period open state in mice.



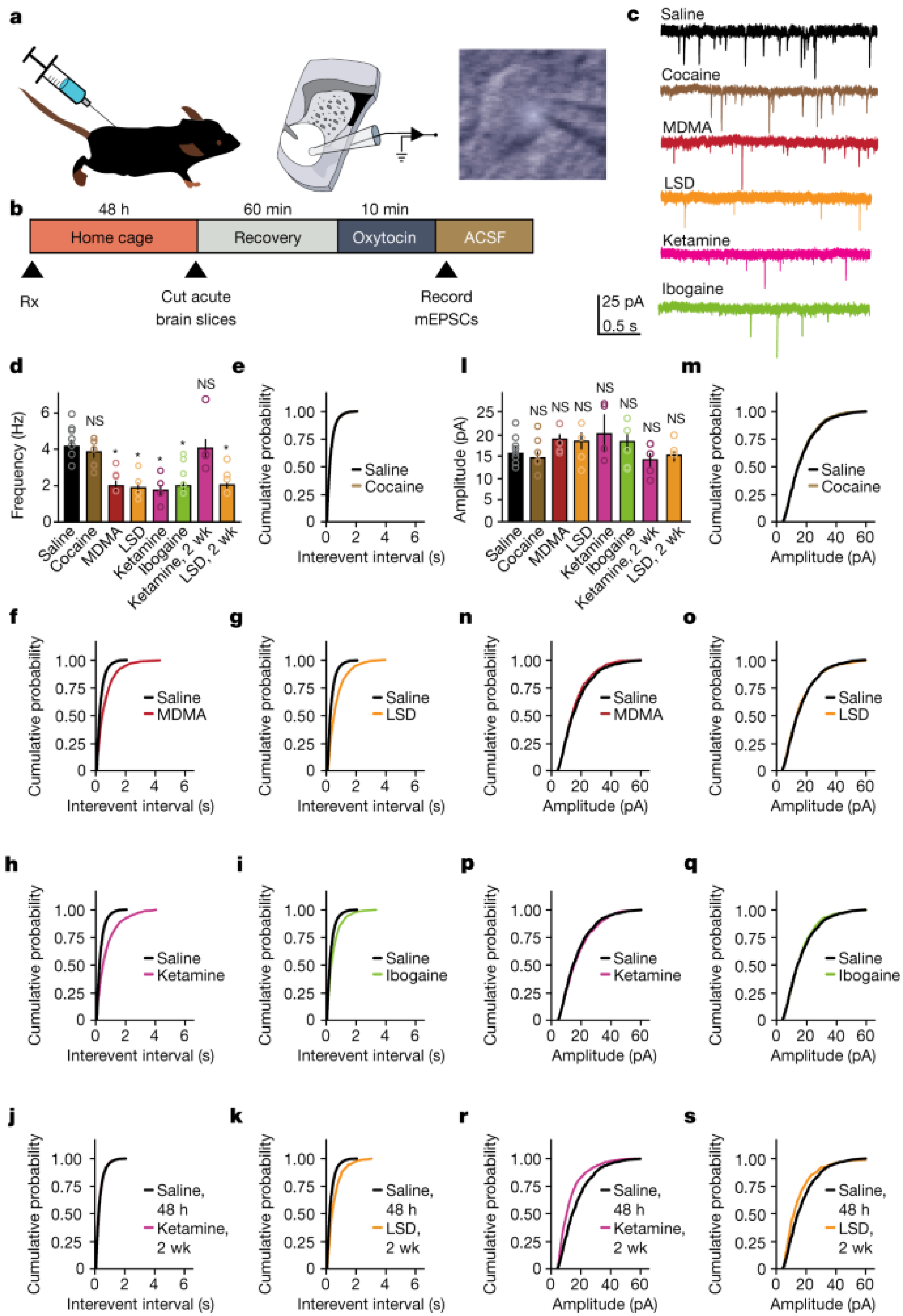
a
Ketamine
(30 min–2 h)
Psilocybin
(3–6 h)
MDMA
(3–6 h)
LSD
(8–10 h)
Ibogaine
(36–72 h)
Acute subjective effects (human)
3
6
9
Time (hours)
b
Ketamine
(48 h)
Psilocybin
(2 weeks)
MDMA
(2 weeks)
LSD
(3 weeks)
Ibogaine
(4 weeks)
Critical period open state (mouse)
1
2
3
4
Time (weeks)

a, Durations of the acute subjective effects of psychedelics in humans (data from refs. [15,16,20,21,22]). **b**, Durations of the critical period open state induced by psychedelics in mice. Based on ref. [11] and Figs. 1 and 2 and Extended Data Fig. 5.

Metaplasticity, not hyperplasticity

Dynamic regulation of the extent to which synaptic plasticity can be induced is called ‘metaplasticity’[23], and is thought to be one of the mechanisms underlying the establishment of critical periods[24]. Previously, we showed that oxytocin induces a novel form of presynaptically expressed long-term depression, and implicated this plasticity in encoding social reward learning[25,26]. Here, to determine whether the ability to induce metaplastic upregulation of oxytocin plasticity generalizes across psychedelics, we pretreated adult mice with either saline, cocaine or psychedelics. Forty-eight hours or two weeks later we prepared ex vivo acute slices containing the nucleus accumbens (NAc) and conducted whole-cell voltage-clamp recordings from medium spiny neurons (MSNs) (Fig. 4a–c). A 10-min bath application of oxytocin induced a significant decrease in the frequency (Fig. 4d–k) but not the amplitude (Fig. 4l–s) of miniature excitatory post-synaptic currents (mEPSCs) following pretreatment with MDMA, LSD, psilocybin, ketamine and ibogaine, but not with saline or cocaine, at 48 h; this metaplasticity persisted for 2 weeks in the LSD pretreatment group, but not in the ketamine pretreatment groups. We did not observe significant changes in baseline mEPSC amplitude or frequency following pretreatment with psychedelics in the NAc or in layer 5 of the medial prefrontal cortex (mPFC) (Extended Data Fig. 6). Together, these results provide evidence that psychedelics induce metaplasticity rather than hyperplasticity, a distinction that is especially important for designing biomarkers to test therapeutic profiles and abuse liability of novel compounds.

Fig. 4: Psychedelics induce metaplasticity.


a
b
48 h
60 min
10 min
Home cage
Recovery
Oxytocin
ACSF
Rx
Cut acute brain slices
Record mEPSCs
c
Saline
Cocaine
MDMA
LSD
Ketamine
Ibogaine
25 pA
0.5 s
d
Frequency (Hz)
NS
*
Saline
Cocaine
MDMA
LSD
Ketamine
Ibogaine
Ketamine, 2 wk
LSD, 2 wk
e
l
Amplitude (pA)
m
Cumulative probability
Interevent interval (s)
Amplitude (pA)
Saline
Cocaine
f
g
n
o
Saline
MDMA
Saline
LSD
h
i
p
q
Saline
Ketamine
Saline
Ibogaine
j
k
r
s
Saline, 48 h
Ketamine, 2 wk
Saline, 48 h
LSD, 2 wk

a,**b**, Illustration (**a**) and time course (**b**) of treatment and electrophysiology protocol. Illustration in **a** adapted from ref. [25]. **c**, Representative mEPSC traces recorded from MSNs in the NAc of oxytocin-treated brain slices collected from mice pretreated with saline ($n = 8$), 20 mg kg^{-1} cocaine ($n = 6$), 10 mg kg^{-1} MDMA ($n = 4$), 1 µg kg^{-1} LSD ($n = 4$), 3 mg kg^{-1} ketamine ($n = 4$) or 40 mg kg^{-1} ibogaine ($n = 5$). **d**–**k**, Average frequency of mEPSCs (**d**) and cumulative probabilities of interevent intervals for cocaine (**e**), MDMA (**f**), LSD (**g**), ketamine (**h**) and ibogaine (**i**) recorded from MSNs after two days, and after two weeks (wk) for ketamine (**j**) and LSD (**k**). **l**–**s**, Average (**l**) and cumulative probability distributions of amplitudes recorded from MSNs for cocaine (**m**), MDMA (**n**), LSD (**o**), ketamine (**p**) and ibogaine (**q**) recorded from MSNs after two days, and after two weeks for ketamine (**r**) and LSD (**s**). One-way analysis of variance revealed a significant effect of treatment on frequency (**d**, $F_{(7,31)} = 5.99$, $P = 0.0002$) but not amplitude (**l**, $F_{(7,31)} = 1.09$, $P = 0.39$), and multiple comparison analysis revealed an oxytocin-mediated decrease in mEPSC frequency after pretreatment with psychedelics (**f**, MDMA: $P = 0.011$; **g**, LSD: $P = 0.0013$; **h**, ketamine: $P = 0.001$; **i**, ibogaine: $P = 0.013$), but not cocaine ($P = 0.83$), and that this decrease remained significant at the two-week time point with LSD (**k**, $n = 4$, $P = 0.01$) but not ketamine (**j**, $n = 4$, $P = 0.99$). All cells have been recorded in slices of adult mice at P98. Data are mean ± s.e.m. *$P < 0.05$; NS, not significant ($P > 0.05$). n refers to the number of biologically independent cells.

5-HT$_{2A}$R is not the universal mechanism

The serotonin receptor 5-HT$_{2A}$R, first identified by its binding to LSD[27], mediates alterations of perception and cognition induced by ‘serotonergic psychedelics’[10] such as LSD[28] and psilocybin[29]. Furthermore, MDMA is thought to trigger synaptic efflux of serotonin through its binding at the serotonin transporter SERT[30], and some of the effects of ketamine are reportedly mediated by 5-HT$_{2A}$R[31]. Thus, we sought to determine the role of 5-HT$_{2A}$R in reopening the social reward learning critical period with LSD, psilocybin, MDMA and ketamine. We administered psychedelics intraperitoneally in P96 adult mice either alone or in combination with

ketanserin (HTR-A, 0.1 mg kg^{-1})—the 5-HT_{2A}R antagonist used in human studies—which we injected 30 min before the psychedelic (Extended Data Fig. 7). Pre-treatment with either LSD or psilocybin induced reinstatement of sCPP measured 48 h later, and this effect was blocked by co-administration of ketanserin (Extended Data Fig. 7). However, MDMA-induced reinstatement of sCPP persisted in the presence of ketanserin (Extended Data Fig. 7). Similarly, co-administration of ketanserin did not block ketamine-induced reinstatement of social reward learning in adulthood (Extended Data Fig. 7). These results demonstrate that whereas 5-HT_{2A}Rs are required for LSD- and psilocybin-induced reopening of the social reward learning critical period (with potential contributions from serotonin 2B and 2C receptors, since ketanserin also has affinity at these serotonin receptor 2 subtypes), MDMA and ketamine reinstate social reward learning in a 5-HT_{2A}R-independent manner. Although some have argued[10,32] that psychedelics that bind 5-HT_{2A}R (such as LSD and psilocybin) should be classified separately from those that do not (such as MDMA and ketamine), these results identify a novel property (critical period reopening) that coheres the category of psychedelics but violates the 5-HT_{2A}R-binding boundary. Thus, combined with the data presented in Figs. 1 and 2, these results support the continued use of the established naming convention for psychedelics[1,2], rather than subclassification or renaming based on receptor binding or subjective properties.

β-arrestin-2 is not the universal mechanism

Recent studies indicate that prolonged binding at the 5-HT_{2A}R by LSD triggers β-arrestin-2 (β-arr2)-biased signalling over canonical G-protein signalling[33]. Moreover, the effects of MDMA[11] and ibogaine[34] are also thought to be mediated by metabotropic G-protein-coupled receptors (GPCRs). Although the therapeutic effects of ketamine are thought to be mediated by ionotropic NMDA receptors[35], the metabotropic glutamate receptor 5 has also been implicated[36]. To test the hypothesis that β-arr2-biased signalling mediates the ability of psychedelics to reopen the social reward learning critical period, we examined their effects in commercially available β-arr2-knockout (KO) mice. We began by determining baseline

sCPP in juvenile and adult β-arr2-KO mice and found that these mice exhibited the normal maturational profile of social reward learning (Extended Data Fig. 8). Next, we compared the magnitude of sCPP in adult (P98) β-arr2 wild-type and β-arr2-KO mice 48 h following administration of psychedelic drugs (Extended Data Fig. 9). LSD and MDMA reopened the social reward learning critical period in wild-type mice but did not do so in β-arr2-KO mice (Extended Data Fig. 9). Conversely, ketamine and ibogaine were able to reinstate social reward learning in both wild-type and β-arr2-KO mice (Extended Data Fig. 9). Together, these results demonstrate that whereas β-arr2 signalling is required for LSD- or MDMA-induced reopening of the social reward learning critical period, ketamine or ibogaine reinstate social reward learning in a β-arr2-independent manner.

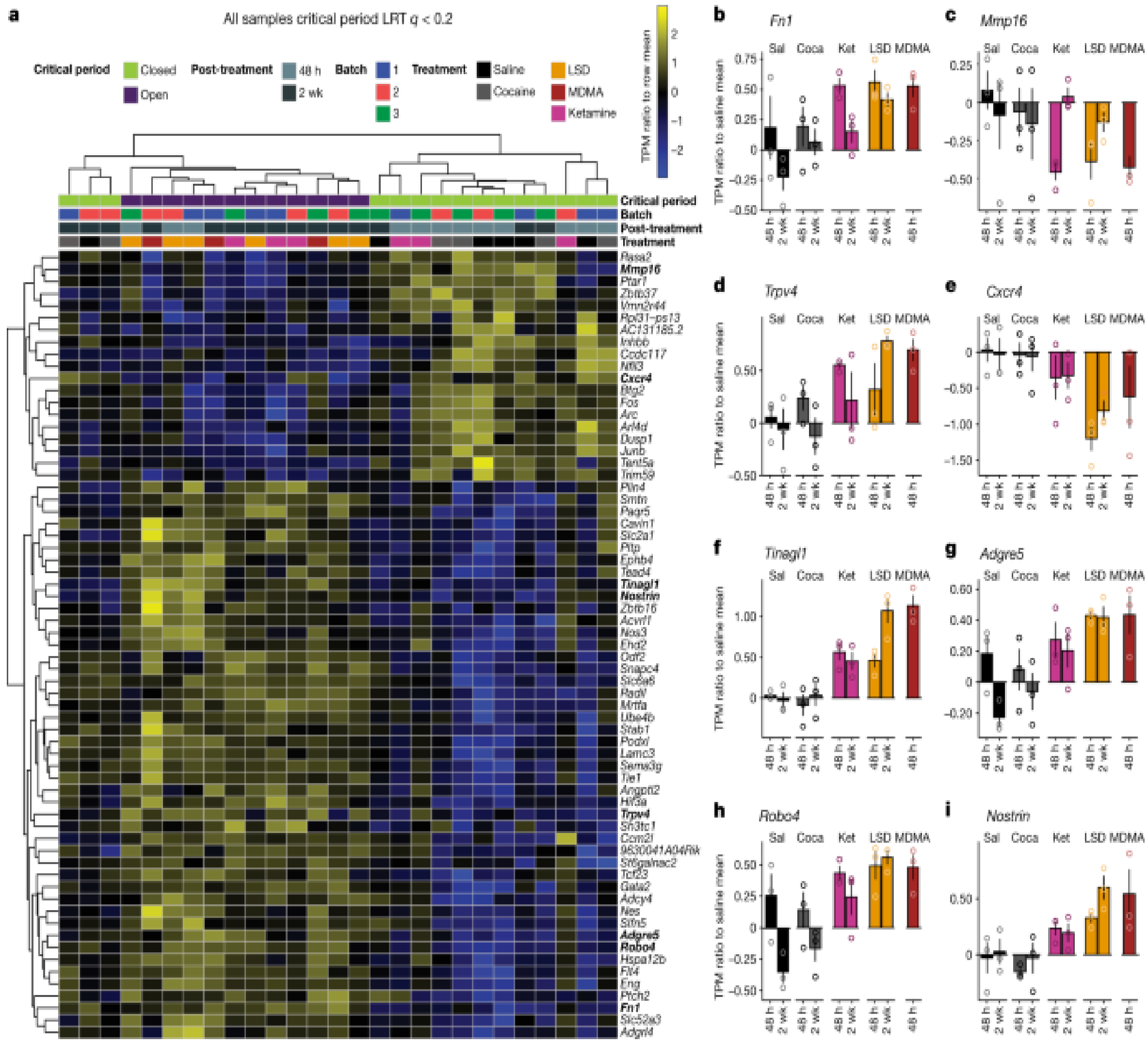
Psychedelics induce remodelling of the ECM

Since psychedelics as a class all reopen the social reward learning critical period (Fig. 1) even though these drugs act on a diverse array of principal binding targets (Extended Data Fig. 7) and biochemical signalling pathways (Extended Data Fig. 9), we reasoned that the common mechanism that enables critical period reopening might be downstream of these cellular processes. Furthermore, given the durability of the response (Fig. 2), we hypothesized that psychedelics may modulate the expression of specific genes or pathways. To test this hypothesis, we carried out RNA sequencing of the microdissected NAc 48 h and 2 weeks after pretreatment with either saline, cocaine, ketamine, LSD or MDMA. We collected total mRNA from each sample and made strand-specific libraries for each of three replicates from each condition. Transcript-level abundances were collapsed to gene-level expression estimates for model fitting.

To directly compare treatment-related transcriptional changes specific to the shared ability of psychedelics to reopen the social reward learning critical period, we analysed the gene expression dataset between conditions in which the critical period is in the open state (48 h and 2 weeks after LSD treatment, 48 h after ketamine treatment, and 48 h after MDMA treatment) versus conditions where the critical period remains in or returns to the closed state (48 h and two weeks after saline treatment, 48 h and two weeks after cocaine treatment, and two weeks after ketamine treatment). Using this

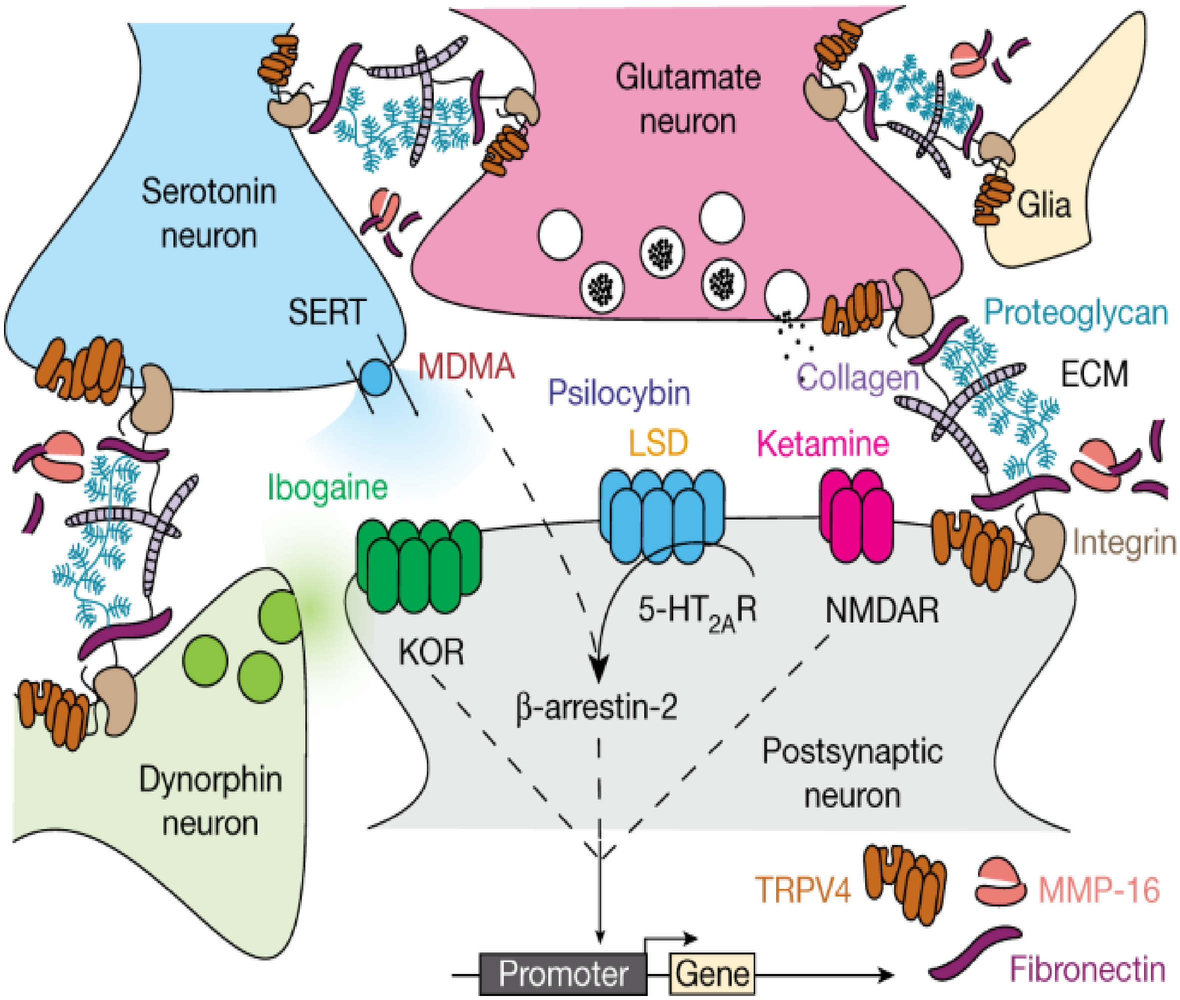
approach, we identified 65 genes that were significantly differentially expressed (likelihood ratio test; Benjamini–Hochberg-corrected $q \leq 0.1$) (Fig. 5). Gene set enrichment analysis of this list identified significant enrichment of ontologies associated with endothelial development, regulation of angiogenesis, vascular development and tissue morphogenesis. Of note, many of the top scoring genes are components of the extracellular matrix (ECM) or have been implicated in its remodelling, including: *Fn1*(ref. [37]), *Mmp16*(ref. [38]), *Trpv4*(ref. [39]), *Tinagl1*(ref. [40]), *Nostrin*[41], *Cxcr4*(ref. [42]), *Adgre5*(ref. [43]), *Robo4*(ref. [44]) and *Sema3g*[45]. Additionally, the differentially expressed gene set includes the immediate early genes (IEGs) *Fos*, *Junb*, *Arc* and *Dusp*. When we did not control for the psychedelic-specific psychoactive response (saline versus all drug conditions, including cocaine), we identified 39 differentially expressed genes (Benjamini–Hochberg-corrected $q \leq 0.15$) (Extended Data Fig. 10); however, enrichment analysis identified no significant ontologies associated with this gene set, and only 6 genes (*Hspa12b*, *Sema3g*, *Eng*, *Flt4*, *Cavin1* and *Ube4b*) overlapped with the differentially expressed genes in the open state versus closed state dataset shown in Fig. 5. These results provide evidence that the shared ability of psychedelics to reopen the social reward learning critical period converges at transcriptional regulation of the ECM. On the basis of these findings, our working model (Fig. 6) posits that psychedelics act at a diverse array of binding targets (such as SERT, 5-HT_{2A}R, NMDA and KOR), to trigger a downstream signalling response that leads to activity-dependent (perhaps via IEG-mediated coincidence detection) degradation of the ECM, which in turn is the permissive event that enables metaplasticity. In this model, transcriptional upregulation of ECM components (for example, FN1) and downregulation of ECM proteolytic enzymes (for example, MMP-16), reflects the homeostatic response to these long-lasting cellular changes. Together, these results demonstrate novel biological effects (behavioural, temporal, electrophysiological and molecular) that—similar to therapeutic effects—are shared across psychedelics.

Fig. 5: Characteristic changes in transcription induced by psychedelics.



a, Heat map of normalized RNA expression values from the microdissected NAc for genes that are significantly differentially expressed in conditions where the critical period remains in the open state versus conditions where the critical period remains in or returns to the closed state. LRT, likelihood ratio test; TPM, transcripts per million. **b**–**i**, Ratio of expression values to average saline baseline for top scoring genes related to extracellular matrix remodelling: *Fn1* (**b**), *Mmp16* (**c**), *Trpv4* (**d**), *Cxcr4* (**e**), *Tinagl1* (**f**), *Adgre5* (**g**), *Robo4* (**h**) and *Nostrin* (**i**). Coca, cocaine; ket, ketamine; sal, saline.

Fig. 6: Working model of convergent cellular mechanisms of psychedelics.



Psychedelics act on a diverse array of principal binding targets and downstream signalling mechanisms that are not limited to the serotonin 2A receptor (Extended Data Fig. 7) or β-arr2 (Extended Data Fig. 9). Instead, mechanistic convergence occurs at the level of DNA transcription (Fig. 5). Dynamically regulated transcripts include components of the extracellular matrix (ECM) such as fibronectin, as well as receptors (such as TRPV4) and proteases (such as MMP-16) implicated in regulating the ECM. Adapted from ref. [25].

Conclusions

These studies provide a novel conceptual framework for understanding the therapeutic effects of psychedelics, which have shown significant promise for treating a wide range of neuropsychiatric diseases, including depression,

PTSD and addiction. Although other studies have shown that psychedelics can attenuate depression-like behaviours[35,46,47,48] and may also have anxiolytic[49], anti-inflammatory[50] and antinociceptive[51] properties, it is unclear how these properties directly relate to the durable and context dependent therapeutic effects of psychedelics[4,6,7,8]. Furthermore, although previous in vitro studies have suggested that psychedelic effects might be mediated by their ability to induce hyperplasticity[52], this account does not distinguish psychedelics from addictive drugs (such as cocaine, amphetamine, opioids, nicotine and alcohol) whose capacity to induce robust, bidirectional, morphological and physiological hyperplasticity is thought to underlie their addictive properties[12]. Moreover, our ex vivo results (Fig. 4 and Extended Data Fig. 6) are consistent with in vivo studies, which demonstrate that dendritic spine formation following administration of psychedelics is both sparse and context dependent[47,53,54], suggesting a metaplastic rather than a hyperplastic mechanism. Indeed, previous studies have also directly implicated metaplasticity in the mechanism of action of ketamine[55,56,57]. At the same time, since our results show that psychedelics do not directly modify addiction-like behaviours (Extended Data Fig. 4 and ref. [11]), they provide a mechanistic clue that critical period reopening may be the neural substrate underlying the ability of psychedelics to induce psychological flexibility and cognitive reappraisal, properties that have been linked to their therapeutic efficacy in the treatment of addiction, anxiety and depression[58,59,60].

Although the current studies have focused on the critical period for social reward learning, critical periods have also been described for a wide variety of other behaviours, including imprinting in snow geese, song learning in finches, language learning in humans, as well as brain circuit rearrangements following sensory or motor perturbations, such as ocular dominance plasticity and post-stroke motor learning[61,62,63,64,65]. Since the ability of psychedelics to reopen the social reward learning critical period is independent of the prosocial character of their acute subjective effects (Fig. 1), it is tempting to speculate that the altered state of consciousness shared by all psychedelics reflects the subjective experience of reopening critical periods. Consistent with this view, the time course of acute subjective effects of psychedelics parallels the duration of the open state induced across compounds (Figs. 2 and 3). Furthermore, since our results point to a shared

molecular mechanism (metaplasticity and regulation of the ECM) (Figs. 4–6) that has also been implicated in the regulation of other critical periods[55,56,57,64,66], these results suggest that psychedelics could serve as a 'master key' for unlocking a broad range of critical periods. Indeed, recent evidence suggests that repeated application of ketamine is able to reopen the critical period for ocular dominance plasticity by targeting the ECM[67,68]. This framework expands the scope of disorders (including autism, stroke, deafness and blindness) that might benefit from treatment with psychedelics; examining this possibility is an obvious priority for future studies.

Methods

Mice

Male wild-type mice were bred in house and weaned at 3 weeks or obtained from Jackson Laboratories (stock no. 000664). β-arr2-KO mice (stock no. 011130) were obtained from Jackson Laboratories, bred in house and weaned at 3 weeks of age. All mice were inbred to the C57BL/6J congenic 'wild-type' strain (as opposed to outbred 'true wilds', which were not used in this study). Congenic strains are generated by backcrossing for a minimum of 10 generations, a standard that is derived from the congenic interval, and the theoretical estimate that by the 10th generation, 99.99% of the congenic strain background will be from the recipient inbred[69]. Although the β-arr2-KO mouse (Jackson Laboratories stock no. 011130), was originally derived on the 129X1/SvJ background[70], it was backcrossed to the C57BL/6J congenic strain at Jackson Laboratories (https://www.jax.org/strain/011130). All mice were maintained on a 12 h:12 h natural light:dark cycle, starting at 07:30 with food and water provided ad libitum. All behavioural experiments were conducted during the same circadian period (07:30–19:30) in a dedicated, sound- and odour-controlled behavioural testing room, which is separated from the vivarium, and no other experiments were conducted simultaneously in the same room. Sample size was estimated based on previous work and published literature. Experimenters were blind to the condition when subjective criteria were used as a component of data analysis, and control and test conditions were interleaved. Mice were randomly assigned to experimental and control

groups. All procedures complied with the animal care standards set forth by the National Institutes of Health and were in accordance with protocols approved by the Johns Hopkins University Animal Care and Use Committee.

sCPP assay

The protocol for sCPP was adapted from previously published work[11]. Mice were socially housed (3–5 males) in a cage containing corncob bedding (Anderson Cob, 0.25 inch cob, Animal Specialties and Provisions) until the pre-determined age for sCPP testing. Each mouse was used for only one behavioural time point. At the pre-determined age, mice were placed in an open field activity chamber (ENV-510, Med Associates) equipped with infrared beams and a software interface (Activity Monitor, Med Associates) to monitor the position of the mouse. The apparatus was partitioned into two equally sized zones using a clear Plexiglas wall, with a 5 cm diameter circular hole at the base; each zone contained one type of novel bedding (Alpha-Dri, Animal Specialties and Provisions or Kaytee Soft Granule, Petco). The amount of time spent freely exploring each zone was recorded during 30-min test sessions. For example, a score of 900 means that the mouse spent exactly 50% of its time on each of the two beddings, whereas a score of 1,800 means that it spent the full 30 min in the bedding that would be subsequently assigned as the social conditioning cue, and no time in the bedding that would be assigned as the isolation conditioning cue. After an initial pre-conditioning trial to establish baseline preference for the two sets of bedding cues, mice were assigned to receive social conditioning (with cage mates) for 24 h on one type of bedding, followed by 24 h of isolation conditioning (without cage mates) on the other bedding cue. To assure unbiased design, chamber assignments were counterbalanced for side and bedding cues. Immediately after the isolation conditioning, a 30-min post-conditioning trial was conducted to establish preference for the two conditioned cues. CPP is a learned association between a condition (for example, social) and a cue (bedding). It does not require scent from the other mice, as the bedding itself serves as the cue. Exclusion criteria for this behaviour are strictly defined as a pre-conditioning preference score of >1.5 or <0.5. Mice are never excluded based on the quality of their social interactions. Pre-conditioning versus post-conditioning social preference

scores were considered significant if paired Student's *t*-test *P* values were less than 0.05. Comparisons between experimental conditions were made using both normalized social preference scores (time spent in social zone post-treatment divided by pre-treatment) and subtracted social preference scores (time spent in social zone post minus pre); these were considered significant if unpaired Student's *t*-test *P* values were <0.05. All experiments were performed during the mouse rest period (light cycle), since pilot experiments revealed that sCPP is most robust if assayed during this period. Prior to i.p. drug treatment experiments (MDMA, LSD, psilocybin, ketamine or ibogaine hydrochloride), mice were habituated to the injection procedure with daily saline i.p. injections in the home cage. Pharmacological delivery schedules were counterbalanced for type of drug. Unless otherwise stated (Fig. 2 and Extended Data Fig. 5), for pretreatment, experiments mice were tested 48 h after the injection to allow for complete clearance of the drug. For the experiment testing involvement of the 5-HT_{2A}R, the 5-HT_{2A}R antagonist ketanserin was administered i.p. 30 min prior to the injection of the drug tested.

Electrophysiology

Subjects received an i.p. injection of either LSD (1 $\mu g\,kg^{-1}$), ketamine (3 $mg\,kg^{-1}$), psilocybin (0.3 $mg\,kg^{-1}$), MDMA (10 $mg\,kg^{-1}$), ibogaine (40 $mg\,kg^{-1}$) or saline. Forty-eight hours after drug treatment, either parasagittal slices containing the NAc core (250 µm thick) or coronal slices containing the PL/IL region of the mPFC (250 µm thick) were prepared from C57BL/6 mice using standard procedures. In brief, after mice were anaesthetized with isoflurane and decapitated, brains were quickly removed and placed in ice-cold low-sodium, high-sucrose dissecting solution (228 mM sucrose, 26 mM $NaHCO_3$, 11 mM glucose, 2.5 mM KCl, 1 mM NaH_2PO_4, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$). Slices were collected with a Leica VT 1200s vibrating microtome. Slices were allowed to recover for a minimum of 60 min in a submerged holding chamber (∼25 °C) containing artificial cerebrospinal fluid (ACSF) consisting of 119 mM NaCl, 2.5 mM KCl, 2.5 mM $CaCl_2$, 1.3 mM $MgCl_2$, 1 mM NaH_2PO_4, 11 mM glucose and 26.2 mM $NaHCO_3$. For hyperplasticity recordings (Extended Data Fig. 6), slices were removed from the holding chamber and placed into the recording chamber, where they

were continuously perfused with oxygenated (95% O_2, 5% CO_2) ACSF at 2 ml min^{-1} at 25 °C. For metaplasticity recordings (Fig. 4), slices were removed from the holding chamber and incubated first for 10 min in oxygenated ACSF containing picrotoxin (50 μM, Sigma), followed by 10-min incubation in oxygenated ACSF containing both picrotoxin and oxytocin (1 μM, Tocris) before being placed into the recording chamber. Whole-cell voltage-clamp recordings from MSNs or layer V pyramidal cells were obtained under visual control using a 40× objective. The NAc core was identified by the presence of the anterior commissure, and the PL/IL region of the mPFC was identified by the presence of the forceps minor of the corpus callosum. Recordings were made with electrodes (2.5–4.0 MΩ) filled with 115 mM $CsMeSO_4$, 20 mM CsCl, 10 mM HEPES, 0.6 mM EGTA, 2.5 mM MgCl, 10 mM sodium phosphocreatine, 4 mM sodium ATP, 0.3 mM sodium GTP and 1 mM QX-314. Miniature EPSCs were collected at a holding potential of −70 mV in the presence of tetrodotoxin (0.5 μM, Tocris Biosciences) and picrotoxin (50 μM, Sigma). Two minutes after break-in, 30-s blocks of events (total of 200 events per cell) were acquired and analysed using the Recording Artist plugin in Igor Pro software with threshold parameters set at 5 pA amplitude and <3 ms rise time. All events included in the final data analysis were verified visually. Data were analysed by multivariate analysis of variance (MANOVA) with three independent variables (drug, brain area and age) and two dependent variables (frequency and amplitude). Likelihood ratio test performed comparing the full model using treatment, age, and structure to a reduced model using age and structure. All calculations were performed in either GraphPad Prism 9 or the R programming language and are available as Supplementary Code 1 and in the repository at https://github.com/genesofeve/DolenPsychedelicOpenState.

RNA extraction and sequencing

Male wild-type C57BL6/J mice were injected i.p. with LSD, ketamine, cocaine (20 mg kg^{-1}) or saline solution either 2 weeks or 48 h before the mice were euthanized. At P98 to P112, mice were euthanized, brains were rapidly removed and ~1mm thick coronal slice ($n = 3$ mice per condition) containing the nucleus accumbens were sectioned using a mouse brain matrix. To microdissect the NAc, slices were placed in a petri dish containing ice-cold ACSF (125 mM NaCl, 2.5 mM KCl, 2 mM $CaCl_2$, 1

mM $MgCl_2$, 1.25 mM NaH_2PO_4, 10 mM glucose and 26 mM $NaHCO_3$) supplemented with RNase inhibitor and oxygenated with carbogen gas (95% O_2 and 5% CO_2) to pH 7.3–7.4. The NAc was identified using the anterior commissure and other structural markers. Between each dissection, blades were replaced and all the instruments and the matrix were cleaned with a solution containing RNase inhibitor. Following dissection, tissue was immediately placed into 0.5 ml Trizol and subjected to a 15 s burst with a tissue homogenizer to lyse the cells. Samples were kept on ice prior to storage at −20 °C. Total RNA were extracted using the RNeasy Kit from Qiagen. The quality of purified RNA was assessed via both a nanodrop and 2100 Bioanalyzer from Agilent. Library preparation was performed using a TruSeq Stranded mRNA kit (Illumina) using the recommended protocol. Individual dual-indexed libraries were quality controlled, pooled, and sequenced on the NovaSeq 6000 platform on a single S1 flowcell to an average depth of 76,841,745 (±8,066,939.82) paired-end 100 bp reads per sample. Reads were pseudoaligned to the mouse GENCODE vM25 (ref. [71]) reference transcriptome using kallisto (v0.46.2) with 100 bootstrapped samples and 6 threads. Defaults were used for all other parameters. Estimated transcript-level abundances were collapsed to gene-level expression estimates and analysed using the sleuth (v0.30.0) R/Bioconductor package. To identify genes with differential expression as a function of samples where the critical period is reopened we performed a likelihood ratio test comparing a full model which included batch, and critical period to a reduced model that only included batch. Time was not used as an explanatory variable in this model fitting. Using this test, we identified 65 genes as significantly differentially expressed at a 10% false discovery rate (Benjamini–Hochberg-corrected $q \leq 0.1$). To identify genes with differential expression as a function of any drug treatment (including cocaine) versus saline we performed a likelihood ratio test comparing a full model that included batch, and 'treated vs untreated' to a reduced model that only included batch. Using this test, we identified 39 genes as significantly differentially expressed at a 15% false discovery rate (Benjamini–Hochberg-corrected $q \leq 0.15$). Time was not used as an explanatory variable in this model fitting. Raw data will be made publicly available (Gene Expression Omnibus accession numbers: GSE230679 and GSM7231202–GSM7231228). Code to reproduce the RNA-seq analysis and associated

figures is provided as Supplementary Code 2 and in the repository at https://github.com/genesofeve/DolenPsychedelicOpenState.

Statistics

All statistical details can be found in the figure legends, including the type of statistical analysis used, *P* values, *n*, degrees of freedom, *t* values and *f* values. Sample sizes were not predetermined by statistical methods; instead they were estimated based on the previously published literature[11]. Data distributions were assumed to be normal. Homogeneity of variance was tested using Levene's test for equality of variances. Comparisons between experimental manipulations were made using a two-tailed Students *t*-test (paired or unpaired, and with or without Welch's correction as appropriate) and MANOVA for comparisons between multiple outcome measures, with $P < 0.05$ considered significant.

Linear, β-spline, loess smoothing and natural spline models evaluated on the previously published time course of normalized social preference scores[11]. Loess smoothing yielded a pseudoinverse at age 41.695 and a knot point of 35 was chosen for both β-spline and natural spline models. The natural spline outperformed the β-spline (adjusted R^2 of 0.1053 versus 0.5554, respectively) with fewer parameters. Residuals were plotted against fitted values and age to check model assumptions. Leave one out cross validation was also used to assess model fit. Control data from all new experiments was used as test data via the predict R function. RSME and R^2 values were comparable between the original model and the new data. Two-way *t*-tests to compare means of controls groups against matched or binned time periods was done to confirm fit to new data. The full model including coefficients for splines, experiment and condition was constructed and tested against reduced models with the final reduced model being reported. MANOVA analysis was carried out using multivariate linear models and the ANOVA function. All statistical comparisons were carried out in the R programming language and can be found in Supplementary Codes 3 and 4 as well as in the repository at https://github.com/genesofeve/DolenPsychedelicOpenState.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw data are publicly available on the repository at https://github.com/genesofeve/DolenPsychedelicOpenState and at the Gene Expression Omnibus (accession numbers GSE230679 and GSM7231202 to GSM7231228).

References

1. Snyder, S. H. *Drugs and the Brain* (W. H. Freeman & Co., 1986).

2. Shulgin, A. T. The background and chemistry of MDMA. *J. Psychoactive Drugs* **18**, 291–304 (1986).

3. Griffiths, R. R. et al. Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: a randomized double-blind trial. *J. Psychopharmacol.* **30**, 1181–1197 (2016).

4. Noller, G. E., Frampton, C. M. & Yazar-Klosinski, B. Ibogaine treatment outcomes for opioid dependence from a twelve-month follow-up observational study. *Am. J. Drug Alcohol Abuse* **44**, 37–46 (2018).

5. Johnson, M. W. & Griffiths, R. R. Potential therapeutic effects of psilocybin. *Neurotherapeutics* **14**, 734–740 (2017).

6. Mitchell, J. M. et al. MDMA-assisted therapy for severe PTSD: a randomized, double-blind, placebo-controlled phase 3 study. *Nat. Med.* **27**, 1025–1033 (2021).

7. Mithoefer, M. C. et al. Durability of improvement in post-traumatic stress disorder symptoms and absence of harmful effects or drug dependency after 3,4-methylenedioxymethamphetamine-assisted

psychotherapy: a prospective long-term follow-up study. *J. Psychopharmacol.* **27**, 28–39 (2013).

8. Carhart-Harris, R. et al. Trial of psilocybin versus escitalopram for depression. *N. Engl. J. Med.* **384**, 1402–1411 (2021).

9. Zarate, C. A. et al. A randomized trial of an *N*-methyl-d-aspartate antagonist in treatment-resistant major depression. *Arch. Gen. Psychiatry* **63**, 856–864 (2006).

10. Nichols, D. E. Psychedelics. *Pharmacol. Rev.* **68**, 264–355 (2016).

11. Nardou, R. et al. Oxytocin-dependent reopening of a social reward learning critical period with MDMA. *Nature* **569**, 116–120 (2019).

12. Russo, S. J. et al. The addicted synapse: mechanisms of synaptic and structural plasticity in nucleus accumbens. *Trends Neurosci.* **33**, 267–276 (2010).

13. Bedi, G., Hyman, D. & De Wit, H. Is ecstasy an 'empathogen'? Effects of 3,4- methylenedioxymethamphetamine on prosocial feelings and identification of emotional states in others. *Biol. Psychiatry* **68**, 1134–1140 (2010).

14. Holze, F. et al. Distinct acute effects of LSD, MDMA, and d-amphetamine in healthy subjects. *Neuropsychopharmacology* **45**, 462–471 (2020).

15. Niesters, M., Martini, C. & Dahan, A. Ketamine for chronic pain: risks and benefits. *Br. J. Clin. Pharmacol.* **77**, 357–367 (2014).

16. MačiIulaitis, R., Kontrimavičiute, V., Bressolle, F. M. M. & Briedis, V. Ibogaine, an anti-addictive drug: pharmacology and time to go further in development. *Hum. Exp. Toxicol.* **27**, 181–194 (2008).

17. Halberstadt, A. L., Chatha, M., Klein, A. K., Wallach, J. & Brandt, S. D. Correlation between the potency of hallucinogens in the mouse head-twitch response assay and their behavioral and subjective effects in other species. *Neuropharmacology* **167**, 107933 (2020).

18. Autry, A. E. et al. NMDA receptor blockade at rest triggers rapid behavioural antidepressant responses. *Nature* **475**, 91–95 (2011).

19. Tzschentke, T. M. Measuring reward with the conditioned place preference (CPP) paradigm: update of the last decade. *Addict. Biol.* **12**, 227–462 (2007).

20. Kolbrich, E. A. et al. Physiological and subjective responses to controlled oral 3,4-methylenedioxymethamphetamine administration. *J. Clin. Psychopharmacol.* **28**, 432–440 (2008).

21. Madsen, M. K. et al. Psychedelic effects of psilocybin correlate with serotonin 2A receptor occupancy and plasma psilocin levels. *Neuropsychopharmacology* **44**, 1328–1334 (2019).

22. Schmid, Y. et al. Acute effects of lysergic acid diethylamide in healthy subjects. *Biol. Psychiatry* **78**, 544–553 (2015).

23. Abraham, W. C. & Bear, M. F. Metaplasticity: the plasticity of synaptic plasticity. *Trends Neurosci.* **19**, 126–130 (1996).

24. Cho, K. K. A., Khibnik, L., Philpot, B. D. & Bear, M. F. The ratio of NR2A/B NMDA receptor subunits determines the qualities of ocular dominance plasticity in visual cortex. *Proc. Natl Acad. Sci. USA* **106**, 5377–5382 (2009).

25. Dölen, G., Darvishzadeh, A., Huang, K. W. & Malenka, R. C. Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature* **501**, 179–184 (2013).

26. Hung, L. W. L. L. W. et al. Gating of social reward by oxytocin in the ventral tegmental area. *Science* **357**, 1406–1411 (2017).

27. Bennett, J. P. & Snyder, S. H. Stereospecific binding of d-lysergic acid diethylamide (LSD) to brain membranes: relationship to serotonin receptors. *Brain Res.* **94**, 523–544 (1975).

28. Preller, K. H. et al. The fabric of meaning and subjective effects in LSD-induced states depend on serotonin 2A receptor activation. *Curr.*

Biol. **27**, 451–457 (2017).

29. Kometer, M., Schmidt, A., Jäncke, L. & Vollenweider, F. X. Activation of serotonin 2A receptors underlies the psilocybin-induced effects on α oscillations, N170 visual-evoked potentials, and visual hallucinations. *J. Neurosci.* **33**, 10544–10551 (2013).

30. Rudnick, G. & Wall, S. C. The molecular mechanism of 'ecstasy' [3,4-methylenedioxy-methamphetamine (MDMA)]: serotonin transporters are targets for MDMA-induced serotonin release. *Proc. Natl Acad. Sci. USA* **89**, 1817–1821 (1992).

31. Murphy, T. J. & Murnane, K. S. The serotonin 2C receptor agonist WAY-163909 attenuates ketamine-induced hypothermia in mice. *Eur. J. Pharmacol.* **842**, 255–261 (2019).

32. McClure-Begley, T. D. & Roth, B. L. The promises and perils of psychedelic pharmacology for psychiatry. *Nat. Rev. Drug Discov.* **21**, 463–473 (2022).

33. Kim, K. et al. Structure of a hallucinogen-activated G_q-coupled 5-HT2A serotonin receptor. *Cell* **182**, 1574–1588.e19 (2020).

34. Maillet, E. L. et al. Noribogaine is a G-protein biased κ-opioid receptor agonist. *Neuropharmacology* **99**, 675–688 (2015).

35. Suzuki, K., Nosyreva, E., Hunt, K. W., Kavalali, E. T. & Monteggia, L. M. Effects of a ketamine metabolite on synaptic NMDAR function. *Nature* **546**, E1–E3 (2017).

36. Esterlis, I. et al. Ketamine-induced reduction in mGluR5 availability is associated with an antidepressant response: an [^{11}C]ABP688 and PET imaging study in depression. *Mol. Psychiatry* **23**, 824–832 (2018).

37. Lau, L. W., Cua, R., Keough, M. B., Haylock-Jacobs, S. & Yong, V. W. Pathophysiology of the brain extracellular matrix: a new target for remyelination. *Nat. Rev. Neurosci.* **14**, 722–729 (2013).

38. Oh, S. Y., Lee, S.-J., Jung, Y. H., Lee, H. J. & Han, H. J. Arachidonic acid promotes skin wound healing through induction of human MSC migration by MT3-MMP-mediated fibronectin degradation. *Cell Death Dis.* **6**, e1750 (2015).

39. Cussac, L.-A. et al. TRPV4 channel mediates adventitial fibroblast activation and adventitial remodeling in pulmonary hypertension. *Am. J. Physiol.* **318**, L135–L146 (2020).

40. Shan, Z.-G. et al. Upregulation of tubulointerstitial nephritis antigen like 1 promotes gastric cancer growth and metastasis by regulating multiple matrix metallopeptidase expression. *J. Gastroenterol. Hepatol.* **36**, 196–203 (2021).

41. Chakraborty, S. & Ain, R. Nitric-oxide synthase trafficking inducer is a pleiotropic regulator of endothelial cell function and signaling. *J. Biol. Chem.* **292**, 6600–6620 (2017).

42. Uckelmann, H. et al. Extracellular matrix protein Matrilin-4 regulates stress-induced HSC proliferation via CXCR4. *J. Exp. Med.* **213**, 1961–1971 (2016).

43. Tjong, W.-Y. & Lin, H.-H. The role of the RGD motif in CD97/ADGRE5-and EMR2/ADGRE2-modulated tumor angiogenesis. *Biochem. Biophys. Res. Commun.* **520**, 243–249 (2019).

44. Cai, H. et al. Roundabout 4 regulates blood-tumor barrier permeability through the modulation of ZO-1, occludin, and claudin-5 expression. *J. Neuropathol. Exp. Neurol.* **74**, 25–37 (2015).

45. Carulli, D., de Winter, F. & Verhaagen, J. Semaphorins in adult nervous system plasticity and disease. *Front. Synaptic Neurosci.* **13**, 672891 (2021).

46. Hesselgrave, N., Troppoli, T. A., Wulff, A. B., Cole, A. B. & Thompson, S. M. Harnessing psilocybin: antidepressant-like behavioral and synaptic actions of psilocybin are independent of 5-HT2R

activation in mice. *Proc. Natl Acad. Sci. USA* **118**, e2022489118 (2021).

47. Moda-Sava, R. N. et al. Sustained rescue of prefrontal circuit dysfunction by antidepressant-induced spine formation. *Science* **364**, eaat8078 (2019).

48. Vargas, M. V. et al. Psychedelics promote neuroplasticity through the activation of intracellular 5-HT2A receptors. *Science* **379**, 700–706 (2023).

49. Young, M. B. et al. Inhibition of serotonin transporters disrupts the enhancement of fear memory extinction by 3,4-methylenedioxymethamphetamine (MDMA). *Psychopharmacology* **234**, 2883–2895 (2017).

50. Flanagan, T. W. & Nichols, C. D. Psychedelics as anti-inflammatory agents. *Int. Rev. Psychiatry* **30**, 363–375 (2018).

51. Schindler, E. A. D. Psychedelics in the treatment of headache and chronic pain disorders. *Curr. Top. Behav. Neurosci.* **56**, 261–285 (2022).

52. Ly, C. et al. Psychedelics promote structural and functional neural plasticity. *Cell Rep.* **23**, 3170–3182 (2018).

53. Shao, L.-X. et al. Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo. *Neuron* **109**, 2535–2544.e4 (2021).

54. Ali, F. et al. Ketamine disinhibits dendrites and enhances calcium signals in prefrontal dendritic spines. *Nat. Commun.* **11**, 72 (2020).

55. Castrén, E. & Monteggia, L. M. Brain-derived neurotrophic factor signaling in depression and antidepressant action. *Biol. Psychiatry* **90**, 128–136 (2021).

56. Crawford, D. C., Ramirez, D. M. O., Trauterman, B., Monteggia, L. M. & Kavalali, E. T. Selective molecular impairment of spontaneous

neurotransmission modulates synaptic efficacy. *Nat. Commun.* **8**, 14436 (2017).

57. Kim, J.-W. et al. Sustained effects of rapidly acting antidepressants require BDNF-dependent MeCP2 phosphorylation. *Nat. Neurosci.* **24**, 1100–1109 (2021).

58. Kuypers, K. P. C. et al. Ayahuasca enhances creative divergent thinking while decreasing conventional convergent thinking. *Psychopharmacology* **233**, 3395–3403 (2016).

59. Davis, A. K., Barrett, F. S. & Griffiths, R. R. Psychological flexibility mediates the relations between acute psychedelic effects and subjective decreases in depression and anxiety. *J. Context. Behav. Sci.* **15**, 39–45 (2020).

60. Agin-Liebes, G. et al. Prospective examination of the therapeutic role of psychological flexibility and cognitive reappraisal in the ceremonial use of ayahuasca. *J. Psychopharmacol.* **36**, 295–308 (2022).

61. Lorenz, K. Der Kumpan in der Umwelt des Vogels. *J. Ornithol.* **83**, 137–213 (1935).

62. Woolley, S. M. N. & Rubel, E. W. Vocal memory and learning in adult Bengalese finches with regenerated hair cells. *J. Neurosci.* **22**, 7774–7787 (2002).

63. Hartshorne, J. K., Tenenbaum, J. B. & Pinker, S. A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition* **177**, 263–277 (2018).

64. Hensch, T. K. & Quinlan, E. M. Critical periods in amblyopia. *Vis. Neurosci.* **35**, E014 (2018).

65. Dromerick, A. W. et al. Critical Period after Stroke Study (CPASS): a phase II clinical trial testing an optimal time for motor recovery after stroke in humans. *Proc. Natl Acad. Sci. USA* **118**, e2026676118 (2021).

66. Fawcett, J. W., Oohashi, T. & Pizzorusso, T. The roles of perineuronal nets and the perinodal extracellular matrix in neuronal function. *Nat. Rev. Neurosci.* **20**, 451–465 (2019).

67. Venturino, A. et al. Microglia enable mature perineuronal nets disassembly upon anesthetic ketamine exposure or 60-Hz light entrainment in the healthy brain. *Cell Rep.* **36**, 109313 (2021).

68. Grieco, S. F. et al. Subanesthetic ketamine reactivates adult cortical plasticity to restore vision from amblyopia. *Curr. Biol.* **30**, 3591–3603.e8 (2020).

69. Berry, M. & Linder, C. in *The Mouse in Biomedical Research* 2nd edn (eds Fox, J. G. et al.) 53–78 (Elsevier, 2007).

70. Bohn, L. M. et al. Enhanced morphine analgesia in mice lacking β-arrestin 2. *Science* **286**, 2495–2498 (1999).

71. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).

Acknowledgements

We thank members of the Dölen laboratory, J. Smith, A. Berghella and N. Family for helpful comments. MDMA was a gift from the National Institute on Drug Abuse (NIDA) and R. Doblin. Psilocybin was a gift from M. Pletnikov. LSD and ibogaine were a gift from NIDA. We are also grateful for research funding from the Klingenstein-Simons Foundation (G.D.), Kavli Neuroscience Discovery Institute Distinguished Postdoctoral Fellowship (G.L.S.-O.), Johns Hopkins Provost Postdoctoral Fellowship (G.L.S.-O.) and the National Institutes of Health: R01MH117127 (G.D.), R01HD098184 (G.D.), R01AG066768 (L.A.G.), R01AG072305 (L.A.G.) and K99NS122085 (G.L.S.-O.).

Author information

Authors and Affiliations

1. The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Romain Nardou, Edward Sawyer, Young Jun Song, Makenzie Wilkinson, Yasmin Padovan-Hernandez, Júnia Lara de Deus, Noelle Wright, Carine Lama, Sehr Faltin, Loyal A. Goff, Genevieve L. Stein-O'Brien & Gül Dölen

2. The Brain Science Institute, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Romain Nardou, Edward Sawyer, Young Jun Song, Makenzie Wilkinson, Júnia Lara de Deus, Noelle Wright, Carine Lama, Sehr Faltin, Genevieve L. Stein-O'Brien & Gül Dölen

3. Department of Oncology, Division of Biostatistics and Bioinformatics, Sidney Kimmel Comprehensive Cancer Center, Baltimore, MD, USA

   Loyal A. Goff

4. McKusick–Nathans Department of Genetic Medicine, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Loyal A. Goff

5. The Kavli Neuroscience Discovery Institute, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Genevieve L. Stein-O'Brien & Gül Dölen

6. The Department of Neurology, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Gül Dölen

7. The Center for Psychedelics and Consciousness Research, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

   Gül Dölen

8. The Wendy Klag Institute for Autism and Developmental Disabilities, Johns Hopkins University, School of Medicine, Baltimore, MD, USA

Gül Dölen

Contributions

G.D. and R.N. designed the study, interpreted results and wrote the paper. R.N., Y.J.S., M.W., C.L., N.W. and S.F. performed and analysed behavioural experiments, E.S., M.W., Y.P.-H. and J.L.d.D. performed electrophysiological experiments, and G.L.S.-O. performed the statistical analysis as well as the RNA sequencing and subsequent analysis. All authors edited the paper.

Corresponding author

Correspondence to Gül Dölen.

Ethics declarations

Competing interests

G.D. is on the scientific advisory board for Compass Pathways, Pangea Botanica and the MIND Foundation.

Peer review

Peer review information

Nature thanks William Wetsel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Social conditioned place preference assay.

Time course (top) and protocol (bottom) for the social conditioned place preference (social CPP) assay used to measure the reward value of social interactions, where an increased amount of time spent in the bedding paired to the socially conditioned cue in the 'post' compared to 'pre' trial is interpreted as evidence for a learned association between the positive reward value of social interactions and a novel bedding cue (social reward learning).

Extended Data Fig. 2 Computation model of social reward critical period.

Natural-spline model curve fitting the normalized time spent in social versus isolate cue following social CPP across 15 postnatal ages of not injected (P21-P112) or saline injected (P126) mice from Nardou *et al*. 2019[11]. Ages were chosen to span major milestones in social and brain development including: weaning (postnatal day 21, P21), onset of puberty (P28), sexual maturity (P42), early adulthood (P60), and mature adulthood (P90). (**a**) Plot showing model fit to training and test data indicating a strong concordance. (**b**) Plot of residuals vs age for ns-spline model. (**C**) Plot of residuals vs fitted values for ns-spline model.

Extended Data Fig. 3 Psychedelic effects are dose, context, and age dependent.

(**a**) Diagram illustrating experimental time course of i.p. pretreatment in social CPP. (**b**,**c**) Individual (top) and average (bottom) responses indicate that adult animals do not develop a preference for the social bedding cue after pretreatment with an anesthetic dose of i.p. ketamine, suggesting an important role for context, since social interactions are precluded at this dose (100 mg/kg) (**b**, n = 18, $t_{(17)} = -1.75$, $P = 0.0986$; **c**, n = 20, $t_{(19)} = -0.89452$, P = 0.382). (**d**) Comparisons (two-tailed unpaired *t*-test) reveal no significant difference in normalized (top) and subtracted (bottom) social preference for saline versus ketamine pretreated animals (**d**, normalized, $t_{(36)} = 0.012694$,

P = 0.990; subtracted, $t_{(36)} = -0.7581$, $P = 0.454$). (**e**) Normalized social preference of adult animals pretreated with saline and ketamine plotted against ns-spline model of the developmental time course of normalized social preference scores. Comparison to the ns-spine model revealed that the magnitude of sCPP in saline (P = 0.9976) and ketamine (P = 0.9921) pretreated animals did not deviate significantly from the "closed" state. (**f**,**g**) Individual (top) and average (bottom) responses indicate that adult animals develop a significant preference for the social bedding cue after pretreatment with 0.1 mg/kg (**b**, n = 11, $t_{(10)} = -2.3288$, $P = 0.0421$) and 0.2 mg/kg psilocybin (**c**, n = 12, $t_{(11)} = -3.5499$, P = 0.0046). (**h**) Comparisons (two-tailed unpaired *t*-test) reveal no significant difference in normalized (top) and subtracted (bottom) social preference for pretreatment with 0.1 mg/kg versus 0.2 mg/kg psilocybin (**h**, normalized, $t_{(21)} = -1.6504$, $P = 0.115$; subtracted, $t_{(21)} = -1.7097$, $P = 0.109$). (**i**) Normalized social preference of adult animals pretreated with 0.1 mg/kg and 0.2 mg/kg psilocybin plotted against ns-spline model of the developmental time course of normalized social preference scores. Comparison to the ns-spine model revealed that the magnitude of sCPP in animals pretreated with 0.1 mg/kg psilocybin (P = 0.7938) did not deviate significantly from the "closed" state, while animals pretreated with 0.2 mg/kg psilocybin (P = 0.04646) demonstrated a significant mean shift in range of the "open" state. (**j**,**k**) Individual (top) and average (bottom) responses indicate that P42 animals develop a significant preference for the social bedding cue after pretreatment with saline (**j**, n = 22, $t_{(21)} = -5.1882$, $P < 0.001$) and MDMA (**k**, n = 22, $t_{(21)} = -5.5432$, P < 0.001) (two-tailed paired *t*-test). (**d**), Comparisons (two-tailed unpaired *t*-test) reveal no significant difference in normalized (top) and subtracted (bottom) social preference for pretreatment with MDMA (10 mg/kg) versus saline (**l**, normalized, $t_{(42)} = 0.6877$, $P = 0.495$; subtracted, $t_{(42)} = 0.29151$, $P = 0.772$). (**m**) Normalized social preference of P42 animals pretreated with saline and MDMA plotted against ns-spline model of the developmental time course of normalized social preference scores. Comparison to the ns-spine model revealed that the magnitude of sCPP in animals pretreated with 10 mg/kg MDMA (P = 0.1831) or saline (P = .7721) did not deviate significantly from the "open" state; comparisons to spline regression model were considered not significant P > 0.1. Data are presented as mean ± s.e.m.

*P < 0.05; n.s., comparisons not significant (P > 0.05). n = X biologically independent animals.

Extended Data Fig. 4 Psychedelic pretreatment does not addiction-like behaviors.

(**a**) Diagram illustrating experimental time course of i.p. injections in cocaine CPP. (**b**,**c**) Individual (top) and average (bottom) time spent in the cocaine paired context indicates that mice pretreated with saline (**b**, n = 8, $t_{(7)} = -4.5102$, $P = 0.0028$) and LSD (**c**, n = 8, $t_{(7)} = -4.8353$, $P = 0.0019$) both develop significantly increased preference for the cocaine context after conditioning with 5 mg/kg cocaine (two-tailed paired *t*-test). (**d**) Comparisons reveal no difference in normalized cocaine preference (top, $t_{(14)} = -1.0051$, $P = 0.332$) and subtracted cocaine preference (bottom, $t_{(14)} = -1.1035$, $P = 0.288$) (two-tailed unpaired *t*-test). (**e**) Locomotor activity during sensitization to amphetamine (20 mg/kg) following 3 days of saline pretreatment. Administration of MDMA after day 12 did not significantly alter the locomotor response to amphetamine (day 12 versus day 16, n = 12, $t_{(11)} = 2.2171$, $P = 0.443$). Data are presented as mean ± s.e.m. *$P < 0.05$; n.s., comparisons not significant ($P > 0.05$). n = X biologically independent animals.

Extended Data Fig. 5 Duration of the open state induced by psychedelics is variable but not extended by increasing dose.

(**a**,**f**,**k**) Diagrams illustrating experimental time course of i.p. pretreatment in social CPP. (**b**,**c**) Individual (top) and average (bottom) responses indicate that adult animals do not develop a significant preference for the social bedding cue 4 days after pretreatment with saline (**b**, n = 8, $t_{(7)} = 0.38628$, $P = 0.711$) and ketamine (3 mg/kg) (**c**, n = 7, $t_{(6)} = -2.2235$, $P = 0.068$) (two-tailed paired *t*-test). (**d**) Comparisons (two-tailed unpaired *t*-test) reveal no significant difference in normalized (top) and subtracted (bottom) social preference for 4 days after pretreatment with ketamine versus saline (**d**, normalized, $t_{(13)} = t = -1.2142$, $P = 0.258$; subtracted, $t_{(13)} = -1.6823$, $P = 0.123$). (**e**) Comparison to the ns-spine model revealed that the magnitude of sCPP did not deviate significantly from the “closed” state four days after

saline (P = 0.508) and ketamine (P = 0.4418). (**g**,**h**) Individual (top) and average (bottom) responses indicate the reinstatement of social CPP is present three weeks after pretreatment with ibogaine (**g**, 40 mg/kg, n = 20, $t_{(19)} = -3.5381$, $P = 0.0022$), but absent with MDMA (**h**, 10 mg/kg, n = 16, $t_{(15)} = -0.11467$, $P = 0.968$) (two-tailed paired *t*-test). (**i**) Comparisons (two-tailed unpaired *t*-test) reveal a significant difference in normalized (top, $t_{(34)} = 2.615$, P = 0.0170) and subtracted (bottom, t(34) = 2.4756, P = 0.0204) social preference 3 weeks after pretreatment with ibogaine versus MDMA. Comparison to the ns-spine model revealed that the magnitude of sCPP demonstrated a significant mean shift in range of the "open" state for 3 weeks after ibogaine (P = 0.01158), while the magnitude did not deviate significantly from the "closed" state for 3 weeks after MDMA (P = 0.9152). (**k**,**l**) Individual (top) and average (bottom) responses indicate that the reinstatement of social CPP is present 48 h after pretreatment with LSD (**k**, n = 6, $t_{(5)} = -3.0103$, $P = 0.0297$), but absent 4 weeks after (**l**, n = 12, $t_{(11)} = -1.1834$, $P = 0.262$). (**m**) Comparisons reveal non-significant difference in normalized (top) and subtracted (bottom) social preference for 48 h versus 4 weeks after pretreatment with LSD (**m**, normalized, $t_{(16)} = 1.2786$, P = 0.2314; subtracted, $t_{(16)} = 1.4209$, P = 0.1792) (two-tailed unpaired t-test). (**n**) Comparison to the ns-spine model revealed that the magnitude of sCPP demonstrated a significant deviation to the "closed" state at 48 h (P = .0915), while the magnitude did not deviate significantly from the "closed" state at 4 weeks (P = 0.3168); comparisons to spline regression model were considered not significant P > 0.1. Data are presented as mean ± s.e.m. *P < 0.05; n.s., comparisons not significant (P > 0.05). n = X biologically independent animals.

Extended Data Fig. 6 Psychedelics do not induce hyperplasticity in the NAc or mPFC.

(**a**) Time course of treatment and electrophysiology protocol. (**b**) Representative mEPSC traces recorded from MSNs in the NAc of brain slices collected from mice pretreated 48 previous with saline (n = 6), 10 mg/kg MDMA (n = 6), 1 μg/kg LSD (n = 4), 3 mg/kg ketamine (n = 4), and 40 mg/kg ibogaine (n = 4). (**c**) Average frequency of mEPSCs and (**d**–**g**) cumulative probabilities of inter event intervals recorded from MSNs. (**h**)

Average and (**i**–**l**) cumulative probability distributions of amplitudes recorded from MSNs. One-way analysis of variance revealed no significant effect of treatment on frequency ($F_{(5,21)} = 1.05$, $P = 0.41$) or amplitude ($F_{(5,21)} = 0.16$, $P = 0.97$). (**m**) Time course of treatment and electrophysiology protocol for mEPSCs recorded in layer 5 pyramidal neurons in the mPFC of brain slices collected from mice pretreated 48 previous with saline (n = 5), 10 mg/kg MDMA (n = 4), 1 µg/kg LSD (n = 6). (**n**) Average frequency of mEPSCs and (**o**–**p**) cumulative probabilities of inter event intervals recorded from layer 5 pyramidal neurons. (**q**) Average and (**r**–**s**) cumulative probability distributions of amplitudes recorded from layer 5 pyramidal neurons. One-way analysis of variance revealed no significant effect of treatment on frequency ($F_{(2,12)} = 0.34$, $P = 0.72$) or amplitude ($F_{(2,12)} = 0.26$, $P = 0.78$). All cells have been recorded in slices of adult animals at P98. Data are presented as mean ± s.e.m. *$P < 0.05$, n.s. comparisons not significant $P > 0.05$. n = X biologically independent cells.

Extended Data Fig. 7 Serotonin 2A receptors are not universally required for critical period reopening.

(**a**) Diagram illustrating experimental time course of i.p. pretreatment in social CPP. (**b**,**c**,**f**,**g**,**j**,**k**,**n**,**o**) Individual (top) and average (bottom) responses of P98 animals indicate that ketanserin (HTR-A) abolished the reopening of social reward learning critical period by LSD (**b**, LSD, n = 9 animals, $t_{(8)} = -4.938$, $P = 0.001$; **c**, LSD + HTR-A, n = 9 animals, $t_{(8)} = 0.210$, $P = 0.839$), and psilocybin (**f**, psilocybin n = 16 animals, $t_{(15)} = -4.494$, $P < 0.001$; **g**, psilocybin + HTR-A, n = 17 animals, $t_{(16)} = -0.515$, $P = 0.613$), but not by MDMA (**j**, MDMA, n = 18 animals, $t_{(17)} = -2.916$, $P = 0.01$; **k**, MDMA + HTR-A, n = 17 animals, $t_{(16)} = -6.737$, $P < 0.001$) and ketamine (**n**, ketamine n = 16 animals, $t_{(15)} = -4.517$, $P < 0.001$; **o**, ketamine + HTR-A, n = 16 animals, $t_{(15)} = -2.952$, $P < 0.001$) (two tailed paired t-test). (**d**,**h**,**l**,**p**) Comparisons of the normalized (top) and subtracted (bottom) social preference between treatment groups reveal a decrease following LSD + HTR-A versus LSD alone (**d**, normalized, $t_{(16)} = 2.427$, $P = 0.027$; subtracted, $t_{(16)} = 2.377$, $P = 0.030$), a decrease following psilocybin + HTR-A vs psilocybin alone (**h**, normalized, $t_{(31)} = 2.114$, $P = 0.043$; subtracted,

$t_{(31)} = 2.475$, $P = 0.019$), but no difference between MDMA and MDMA + HTR-A pretreatment groups (**l**, normalized, $t_{(33)} = -0.971$, $P = 0.339$; subtracted, $t_{(33)} = -1.282$, $P = 0.209$), nor ketamine and ketamine + HTR-A pretreatment groups (**p**, normalized, $t_{(30)} = 0.013$, $P = 0.990$; subtracted, $t_{(30)} = 0.535$, $P = 0.597$) (two tailed unpaired t-test). (**e**,**i**,**m**,**q**) Normalized social preference in mice pretreated with LSD (**e**), psilocybin (**i**), MDMA (**m**) and ketamine (**q**) in the presence or the absence of HTR-A plotted against ns-spline model of the developmental time course of normalized social preference scores. Comparison to the ns-spine model revealed that the magnitude of sCPP did not deviate significantly from the "closed" state for LSD + HTR-A ($P = 0.728$), and psilocybin + HTR-A ($P = 0.987$), while the magnitude demonstrated a significant mean shift in range of the "open" state for LSD ($P = 2.52e\text{-}09$), psilocybin ($P = 1.43e\text{-}06$), MDMA in presence ($P = 1.77e\text{-}05$) or absence of HTR-A ($P = 0.22e\text{-}4$), and ketamine in presence ($P = 0.002619$) or absence of HTR-A ($P = 0.000996$); comparisons to spline regression model were considered not significant $P > 0.1$. Data are presented as mean ± s.e.m. *$P < 0.05$, n.s. comparisons not significant $P > 0.05$. n = X biologically independent animals.

Extended Data Fig. 8 β-arrestin 2 KO mice exhibit normal maturational profile of social reward learning.

(**a**) Diagram illustrating experimental time course of social CPP. (**b**–**e**) Social CPP in juvenile (P42) and adult (P98) β-arrestin 2 KO mice. (**b**,**c**) Individual (top) and average (bottom) responses of adult animals indicate that only juvenile β-arrestin 2 KO mice develop a preference for the social bedding cue (P42 (n = 17 animals, $t_{(16)} = -4.392$, $P < 0.001$), P98 (n = 17 animals, $t_{(16)} = -0.922$, $P = 0.370$) (two tailed paired t-test)). (**d**) Comparisons reveal a difference in normalized (top, $t_{(32)} = 2.110$, $P = 0.043$) and subtracted (bottom, $t_{(32)} = 2.120$, $P = 0.042$) social preference at P42 versus P98 mice (two tailed unpaired t-test). (**e**) Normalized social preference of β-arrestin 2 KO mice at P42 and P98 plotted against ns-spline model of the developmental time course of normalized social preference scores. Comparison to the ns-spine model revealed that the magnitude of sCPP did not deviate significantly from the "open" state at P42 ($P = 0.840$) and from the "closed" state at P98 ($P = 0.760$); comparisons to spline

regression model were considered not significant $P > 0.1$. Data are presented as mean ± s.e.m. *$P < 0.05$, n.s. comparisons not significant $P > 0.05$. n = X biologically independent animals.

Extended Data Fig. 9 β-arrestin 2 is not universally required for critical period reopening.

(**a**) Diagram illustrating experimental time course of i.p. pretreatment in social CPP. (**b**,**c**,**f**,**g**,**j**,**k**,**n**,**o**) Individual (top) and average (bottom) responses of P98 animals indicate that LSD (**b**, WT, n = 21 animals, $t_{(20)} = -3.992$, $P < 0.001$; **c**, β-arrestin 2 KO, n = 17 animals, $t_{(16)} = -1.105$, $P = 0.286$) and MDMA (**f**, WT, n = 15 animals, $t_{(14)} = -3.994$, $P = 0.001$; **g**, β-arrestin 2 KO, n = 18 animals, $t_{(17)} = -0.759$, $P = 0.458$) don't reopen the social reward learning critical period in β-arrestin 2 KO mice whereas ketamine (**j**, WT, n = 10 animals, $t_{(9)} = -3.448$, $P = 0.007$; **k**, β-arrestin 2 KO, n = 12 animals, $t_{(11)} = -2.903$, $P = 0.014$) and ibogaine (**n**, WT, n = 16 animals, $t_{(15)} = -3.135$, $P = 0.007$; **o**, β-arrestin 2 KO, n = 18 animals, $t_{(17)} = -2.655$, $P = 0.016$) do reopen the critical period both in WT and in β-arrestin 2 KO mice (two tailed paired t-test). (**d**,**h**,**l**,**p**) Comparisons of the normalized (top) and subtracted (bottom) social preference in WT and β-arrestin 2 KO mice reveal a difference in the magnitude of sCPP after pretreatment with LSD (**d**, normalized, $t_{(36)} = 2.248$, $P = 0.031$; subtracted, $t_{(36)} = 2.139$, $P = 0.039$), and MDMA (**h**, normalized, $t_{(31)} = 2.227$, $P = 0.033$; subtracted, $t_{(31)} = 2.112$, $P = 0.043$), but no difference after ketamine (**l**, normalized ($t_{(20)} = -0.545$, $P = 0.591$, subtracted, $t_{(20)} = -0.676$, $P = 0.507$) and ibogaine (**p**, normalized, $t_{(33)} = 0.790$, $P = 0.435$; subtracted, $t_{(33)} = 0.462$, $P = 0.647$) (two tailed unpaired t-test). (**e**,**i**,**m**,**q**) Normalized social preference in WT and β-arrestin 2 KO mice pretreated with LSD (**e**), MDMA (**i**), ketamine (**l**), and ibogaine (**q**) plotted against ns-spline model of the developmental time course of normalized social preference scores of male mice. Comparison to the ns-spine model revealed that the magnitude of sCPP did not deviate significantly from the "closed" state in β-arrestin 2 KO mice for LSD ($P = 0.357$) and MDMA ($P = 0.490$), while the magnitude demonstrated a significant mean shift in range of the "open" state for WT pretreated with LSD (WT, $P = .00315$), or MDMA (WT, $P = .0140$), and for WT and KO

pretreated with ketamine (WT, $P = .00911$; β-arrestin 2 KO, $P = .0198$) or ibogaine (WT, $P = .0249$; β-arrestin 2 KO, $P = .0954$); comparisons to spline regression model were considered not significant $P > 0.1$. Data are presented as mean ± s.e.m. *$P < 0.05$, n.s. comparisons not significant $P > 0.05$. n = X biologically independent animals.

Extended Data Fig. 10 Comparison of gene expression for saline versus all psychoactive drug (including cocaine) treatment groups.

(**a**) Heatmap of normalized RNA expression values from the microdissected NAc for genes significantly differentially expressed between treatment with any psychoactive drug including cocaine vs. the saline control. (**b**–**i**) Ratio of expression values to average saline baseline for top scoring genes from this analysis related to increased synaptic transmission.

Supplementary information

Reporting Summary

Supplementary Code

Code to reproduce all the analyses in this Article. This version was frozen at time of resubmission. Subsequent versions, reformatted for increased accessibility, can be found at https://github.com/genesofeve/DolenPsychedelicOpenState. This project is licensed under the terms of the MIT license. Source Code 1 (MANOVA analysis relating to Fig. 4 and Extended Data Fig. 6), Source Code 2 (RNA sequencing analysis relating to Fig. 5 and Extended Data Fig. 10), Source Code 3 and Source Code 4 (regression analysis for Figs. 1 and 2 and Extended Data Figs. 2, 3, 5 and 7–9).

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation,

distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Nardou, R., Sawyer, E., Song, Y.J. *et al.* Psychedelics reopen the social reward learning critical period. *Nature* **618**, 790–798 (2023). https://doi.org/10.1038/s41586-023-06204-3

- Received: 31 August 2021
- Accepted: 11 May 2023
- Published: 14 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06204-3

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

How psychedelic drugs achieve their potent health benefits

- Sara Reardon

Nature News 14 Jun 2023

Timing is key for behavioural benefits of psychedelics

- Samuel C. Woodburn
- Alex C. Kwan

Nature News & Views 14 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06204-3

- Article
- Open Access
- Published: 14 June 2023

A tripartite rheostat controls self-regulated host plant resistance to insects

- Jianping Guo,
- Huiying Wang,
- Wei Guan,
- Qin Guo,
- Jing Wang,
- Jing Yang,
- Yaxin Peng,
- Junhan Shan,
- Mingyang Gao,
- Shaojie Shi,
- Xinxin Shangguan,
- Bingfang Liu,
- Shengli Jing,
- Jing Zhang,
- Chunxue Xu,
- Jin Huang,
- Weiwei Rao,
- Xiaohong Zheng,
- Di Wu,
- Cong Zhou,
- Bo Du,
- Rongzhi Chen,
- Lili Zhu,
- Yuxian Zhu,

- Linda L. Walling,
- Qifa Zhang &
- …
- Guangcun He

Show authors

Nature **volume 618**, pages 799–807 (2023)

- 8982 Accesses
- 53 Altmetric
- Metrics details

Abstract

Plants deploy receptor-like kinases and nucleotide-binding leucine-rich repeat receptors to confer host plant resistance (HPR) to herbivores[1]. These gene-for-gene interactions between insects and their hosts have been proposed for more than 50 years[2]. However, the molecular and cellular mechanisms that underlie HPR have been elusive, as the identity and sensing mechanisms of insect avirulence effectors have remained unknown. Here we identify an insect salivary protein perceived by a plant immune receptor. The BPH14-interacting salivary protein (BISP) from the brown planthopper (*Nilaparvata lugens* Stål) is secreted into rice (*Oryza sativa*) during feeding. In susceptible plants, BISP targets *O. satvia* RLCK185 (*Os*RLCK185; hereafter *Os* is used to denote *O. satvia*-related proteins or genes) to suppress basal defences. In resistant plants, the nucleotide-binding leucine-rich repeat receptor BPH14 directly binds BISP to activate HPR. Constitutive activation of *Bph14*-mediated immunity is detrimental to plant growth and productivity. The fine-tuning of *Bph14*-mediated HPR is achieved through direct binding of BISP and BPH14 to the selective autophagy cargo receptor *Os*NBR1, which delivers BISP to *Os*ATG8 for degradation. Autophagy therefore controls BISP levels. In *Bph14* plants, autophagy restores cellular homeostasis by downregulating HPR when feeding by brown planthoppers ceases. We identify an insect saliva protein

sensed by a plant immune receptor and discover a three-way interaction system that offers opportunities for developing high-yield, insect-resistant crops.

Main

During the millions of years of plant and insect co-evolution, herbivorous insects have developed diverse feeding strategies to retrieve nutrients from host plants. Such insects can cause up to 18% of annual losses of global crop yield[3]. Because plants are sessile and cannot escape continuous attack by pests, plants have evolved various defence strategies to combat herbivory. Unlike microbial pathogens, insects can move swiftly on or among plants as they feed, making it challenging to understand the mechanisms that underlie their interactions.

Herbivorous insects actively select their feeding sites and secrete saliva during feeding to facilitate nutrient acquisition from host plants[1,4,5]. Some salivary proteins induce host defences[1,5]. In the past decade, plant genes that encode nucleotide-binding leucine-rich repeat (NLR) receptors that confer resistance to piercing and sucking insects have been isolated through map-based cloning[6,7,8,9,10,11,12]. NLR receptors function as intracellular immune receptors that directly or indirectly detect cognate effectors[13,14] and activate effector-triggered immunity (ETI)[6,7,8,9,10], also known as HPR. The identities of the insect avirulence effectors that trigger ETI have not yet been discovered. Hence, how a host plant recognizes the insect and deploys and modulates resistance remains largely unknown.

The brown planthopper (BPH; *N. lugens* Stål) is the most destructive insect pest that threatens rice production globally[15]. Using its stylet mouthparts to probe and penetrate plant cells, it consumes phloem sap and eventually causes plant death and substantial yield losses[16,17,18]. BPHs migrate long distances, crossing national boundaries in search of better host plants[15,16,17,18], which makes regional control of BPH challenging. For this reason, HPR has been crucial for BPH control. The first BPH-resistance gene isolated in rice, *Bph14*, encodes a coiled-coiled nucleotide-binding leucine-rich repeat (CC-NB-LRR) protein that forms a homomeric complex

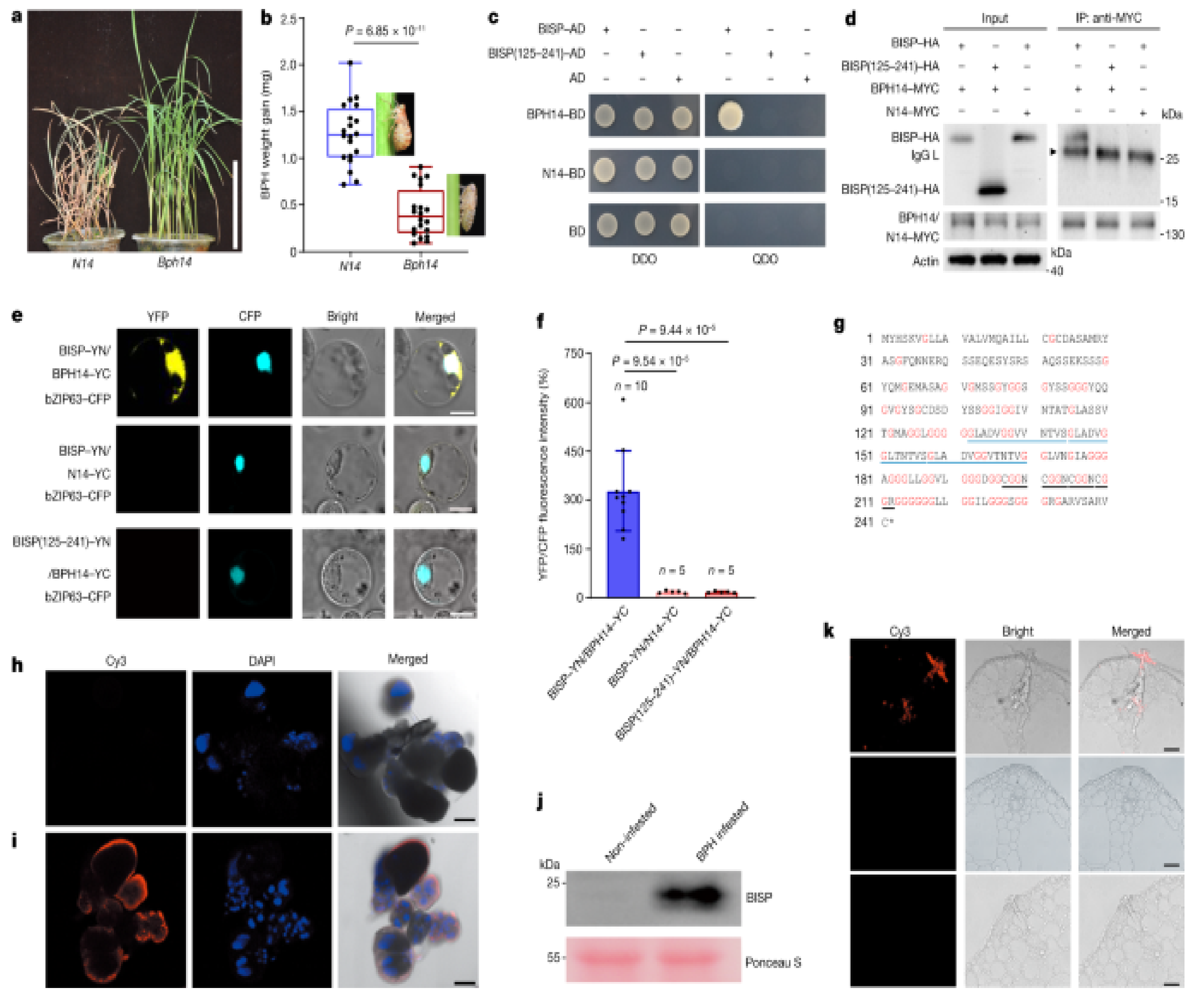
and interacts with WRKY transcription factors[19]. BPH14 enhances WRKY activity to confer resistance to BPH[19].

Here we show that the BPH salivary protein BISP is directly recognized by BPH14. Notably, the perception of BISP not only triggers *Bph14*-mediated resistance but also activates NBR1-mediated selective autophagy. This results in the degradation of BISP to restrict hyperactivation of *Bph14*-regulated defences. We elucidate the molecular and cellular interactions that occur after BPH14 binds to the BISP effector to activate and modulate *Bph14*-mediated HPR.

BPH BISP interacts with rice BPH14

Bph14 has been widely used in rice breeding programmes, and many *Bph14*-resistant varieties have been released for rice production (Supplementary Table 1). Rice plants carrying *Bph14* show little damage, whereas susceptible *N14* plants die 7 days after BPH infestation (Fig. 1a and Extended Data Fig. 1a). Insect feeding and growth were inhibited on *Bph14* plants, which resulted in lower weight gain and honeydew excretion than insects feeding on *N14* plants expressing the susceptible BPH14 protein form N14 (Fig. 1b and Extended Data Fig. 1b). N14 shares 83% sequence identity with BPH14 (ref. [8]). As *Bph14* encodes a typical NLR receptor[8,19], a BPH effector protein may be recognized by BPH14 to activate *Bph14*-mediated resistance.

Fig. 1: The BPH salivary protein BISP interacts with BPH14 and is delivered into rice.



a, Phenotypes of rice plants carrying *Bph14* and *N14* after 7 days of BPH infestation. **b**, Weight gain and representative images of BPH insects feeding on *Bph14* and *N14* plants for 2 days ($n = 20$, biologically independent samples). The box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data and the centre line indicates the median. Individual data points are plotted. **c**–**e**, BISP interacted with BPH14 in Y2H (**c**), co-IP (**d**) and BiFC (**e**) assays. N14 and BISP(125–241) served as negative controls. DDO, SD/-Leu-Trp; QDO, SD/-Leu-Trp-His-Ade. bZIP63–CFP, nuclear marker. **f**, Quantification of relative YFP intensities in BiFC assays. Data are the mean ± s.d. (n = numbers of biologically independent cells). **g**, Amino acid sequence of BISP. The asterisk indicates the stop codon. Glycine residues are marked in red. The 13-amino acid and 4-amino acid tandem repeats are underlined in blue and black, respectively. **h**,**i**, Immunohistochemical localization of BISP in female BPH salivary glands using pre-immune rabbit serum (**h**) or anti-BISP antibodies (**i**). Red

fluorescence (Cy3) and blue fluorescence show the localization of BISP and DAPI-stained nuclei, respectively. **j**, BISP was delivered into rice leaf sheaths during BPH infestation. Leaf sheath proteins were analysed using anti-BISP antibodies. Ponceau S staining served as the loading control. **k**, Immunohistochemical staining showing BISP in BPH-infested rice leaf sheaths. The non-infected (middle) and BPH-infested (bottom) sheaths were detected by anti-BISP antibodies and pre-immune rabbit serum, respectively, served as negative controls. In **b** and **f**, *P* values were derived by one-way analysis of variance (ANOVA). Experiments (**a**–**e**,**h**–**k**) were repeated at least three times, each giving similar results. The results of the other two repeats are presented in Supplementary Fig. 2. Scale bars, 5 μm (**e**), 25 μm (**k**), 100 μm (**h**,**i**) or 10 cm (**a**).

Source data

Planthoppers secrete salivary proteins into rice plants that may induce or suppress host defence responses during feeding[1,4,5,16]. Yeast two-hybrid (Y2H) screens identified BPH-secreted proteins (effectors) that interacted with BPH14. Of the 12 genes identified, one encoded the *Bisp* transcript that was abundant in the transcriptome of the BPH salivary gland (Supplementary Table 2). BISP interacted with BPH14, but not N14, based on repeated Y2H and co-immunoprecipitation (co-IP) experiments (Fig. 1c,d). Expression of BISP tagged with green fluorescent protein (BISP–GFP) in rice protoplasts showed that it localized to the nucleocytoplasm (Extended Data Fig. 1c,d), which is consistent with the subcellular localization of BPH14 (ref. [19]). In addition, bimolecular fluorescence complementation (BiFC) confirmed the interaction between BISP and BPH14 and their nucleocytoplasmic co-localization in rice protoplasts (Fig. 1e,f and Extended Data Fig. 1e).

Bisp (LOC111051577) resides on chromosome 12 of BPH and encodes a 241-amino acid protein with an amino-terminal signal peptide (amino acids 1–25) and no transmembrane domain (Fig. 1g), which suggests that BISP is a secretory protein. Rich in glycine residues (32%), BISP has three 13-amino acid (GLADVGGLTNTVS) and 4-amino acid (CGGN/R) tandem repeats in its carboxy-terminal region (Fig. 1g). Quantitative PCR with reverse transcription (RT–qPCR) confirmed that *Bisp* is highly expressed in salivary glands and in whole female adults (Extended Data Fig. 1f,g).

Immunohistochemical analysis using BISP antiserum revealed that BISP accumulated to higher levels in female salivary glands than in male salivary glands and guts (Fig. 1h,i and Extended Data Fig. 1h–m,p), a result that was consistent with *Bisp* RNA levels. Moreover, BISP signals were higher than the salivary protein NlSP1 (Extended Data Fig. 1n–p).

BPHs were allowed to feed on rice plants to determine whether BISP is secreted into rice leaf sheaths. Immunoblots detected BISP in leaf sheath protein extracts from BPH-infested but not non-infested plants (Fig. 1j). Immunolocalization assays verified the delivery of BISP into rice leaves. BISP was detected along the penetration path of stylets within the leaf sheaths (Fig. 1k). Immunogold electron microscopy further confirmed that BISP was secreted into rice tissues during BPH feeding (Extended Data Fig. 1q–t).

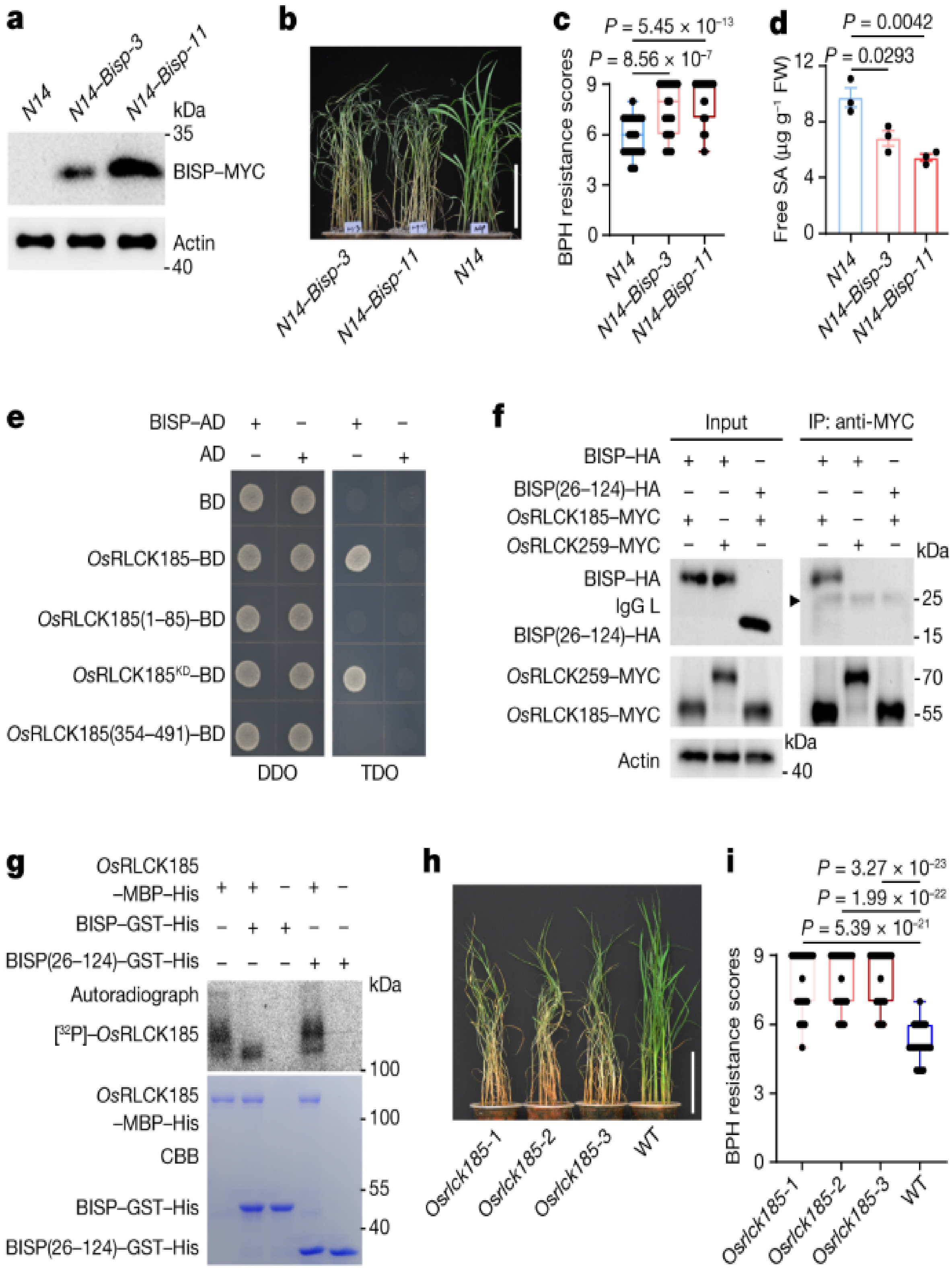
BISP suppresses plant defence

To determine whether BISP influenced the success of BPH on rice, we disrupted *Bisp* expression by microinjecting *Bisp* double-stranded RNAs (dsRNAs) into individual BPHs to induce RNA-mediated interference (RNAi). *Bisp* transcripts and proteins were significantly reduced in BPHs that received *Bisp*-RNAi compared with those that received *GFP*-RNAi (Extended Data Fig. 2a). *Bisp*-RNAi insects exhibited lower weight gain and honeydew excretion (Extended Data Fig. 2b,c). These insects also experienced significantly higher mortality when fed on the susceptible *N14* rice plants than BPHs that received no injection (controls) or *GFP*-RNAi (Extended Data Fig. 2d). Therefore, *Bisp* plays a crucial role in the feeding and performance of BPHs on susceptible rice plants.

To determine the function of BISP in rice, we ectopically expressed BISP (without its signal peptide) in BPH-susceptible *N14* rice (*N14–Bisp*) plants (Fig. 2a). The stature of *N14–Bisp* plants was similar to control *N14* plants (Extended Data Fig. 3a,b). However, *N14–Bisp* plants were hypersensitive to BPHs, displayed more severe symptoms after BPH feeding and died more quickly than the BPH-infested *N14* plants (Fig. 2b,c). Moreover, the weight gain and honeydew excretion of BPHs feeding on *N14–Bisp* plants were significantly higher than those feeding on *N14* plants (Extended Data Fig.

3c,d). A two-host choice test showed that BPHs preferentially settled on *N14–Bisp* plants rather than *N14* plants (Extended Data Fig. 3e). These results show that BISP increases the susceptibility of rice to BPHs.

Fig. 2: BISP interacts with *Os*RLCK185 and suppresses rice defence responses.



a, Immunoblot detection of BISP in *N14* and *N14–Bisp* transgenic rice lines. **b**,**c**, Phenotypes (**b**) and BPH resistance scores (**c**) of *N14* and *N14–Bisp*

plants after 4 days of BPH infestation. *n* = 36 plants examined over 3 independent experiments. **d**, Endogenous free SA levels in *N14–Bisp* and *N14* plants. FW, fresh weight. Data are the mean ± s.d. (*n* = 3, biologically independent experiments). **e**, BISP interacted with *Os*RLCK185 and its kinase domain in a Y2H assay. *Os*RLCK185(1–85), *Os*RLCK185^{KD}, *Os*RLCK185(354–491), *Os*RLCK185 deletion mutants; KD, kinase domain (amino acids 86–353); TDO, SD/-Leu-Trp-His with 1.5 mM 3-AT (3-amino-1,2,4-triazole). **f**, Co-IP showing the interaction between BISP and *Os*RLCK185. BISP(26–124) and *Os*RLCK159 served as negative controls. **g**, *Os*RLCK185 autophosphorylation activity was reduced by BISP. BISP(26–124)–GST–His proteins served as the negative control. Kinase activity was detected by autoradiography and input proteins shown after Coomassie Brilliant Blue staining (CBB). **h**,**i**, Phenotypes (**h**) and BPH-resistance scores (**i**) of Os*rlck185* and WT plants after 4 days of BPH infestation. *n* = 45 plants examined over 3 independent experiments. In box plots in **c** and **i**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data and the centre line indicates the median. Individual data points are plotted. In **c**,**d** and **i**, *P* values were derived by one-way ANOVA. The experiments (**a**,**b**,**d**–**h**) were repeated at least three times, each giving similar results. Scale bars, 10 cm (**b**,**h**).

Source data

We reasoned that the BISP effector was acting as a virulence factor that suppresses rice defence responses, which makes susceptible plants more vulnerable to BPHs. The endogenous levels of free salicylic acid (SA), a plant hormone vital for BPH resistance in rice[10,11,19], were lower in *N14–Bisp* plants than in *N14* plants (Fig. 2d). Similarly, the levels of SA biosynthesis and signalling genes (Os*ICS1* and Os*NPR1*) and downstream defence-related genes (Os*PR1a*, Os*PR1b*, Os*PR5* and Os*PR10*) transcripts were all significantly lower in *N14–Bisp* plants than in *N14* plants (Extended Data Fig. 3f).

Receptor-like kinases (RLKs) play crucial roles in plant defence against pathogens and herbivores[1,20]. We aimed to identify the target of BISP and to elucidate the molecular mechanism that underlies its ability to suppress defence responses in susceptible plants. To that end, we examined the interactions of RLKs, receptor-like cytoplasmic kinases (RLCKs) and MAP

kinases with BISP in Y2H assays (Fig. 2e and Extended Data Fig. 4a–c). Among the eight kinases tested, only *Os*RLCK185 interacted with BISP. BISP interacted with the *Os*RLCK185 kinase domain (86–353 amino acids). The association of *Os*RLCK185 with BISP was confirmed by co-IP in rice protoplasts (Fig. 2f). As *Os*RLCK185 regulates plant immunity through autophosphorylation[21,22], an in vitro phosphorylation assay and immunoblotting with phosphoserine/phosphothreonine-specific antibodies confirmed that *Os*RLCK185 was an active kinase with autophosphorylation activity (Extended Data Fig. 4d). Furthermore, *Os*RLCK185 autophosphorylation was attenuated by BISP (Fig. 2g and Extended Data Fig. 4d).

We used three *Os*RLCK185 knockout rice lines (Os*rlck185*-1–Os*rlck185*-3 plants) to further explore the roles of *Os*RLCK185 in rice defence against BPHs (Extended Data Fig. 4e). Os*rlck185* lines were more susceptible to BPHs, as they were severely damaged compared with wild-type (WT) plants 4 days after BPH infestation (Fig. 2h,i). BPHs that fed on the Os*rlck185* lines gained significantly more weight and excreted significantly more honeydew than those fed on WT plants (Extended Data Fig. 4f,g), which indicated that *Os*RLCK185 positively regulates basal immunity in BPH-susceptible plants.

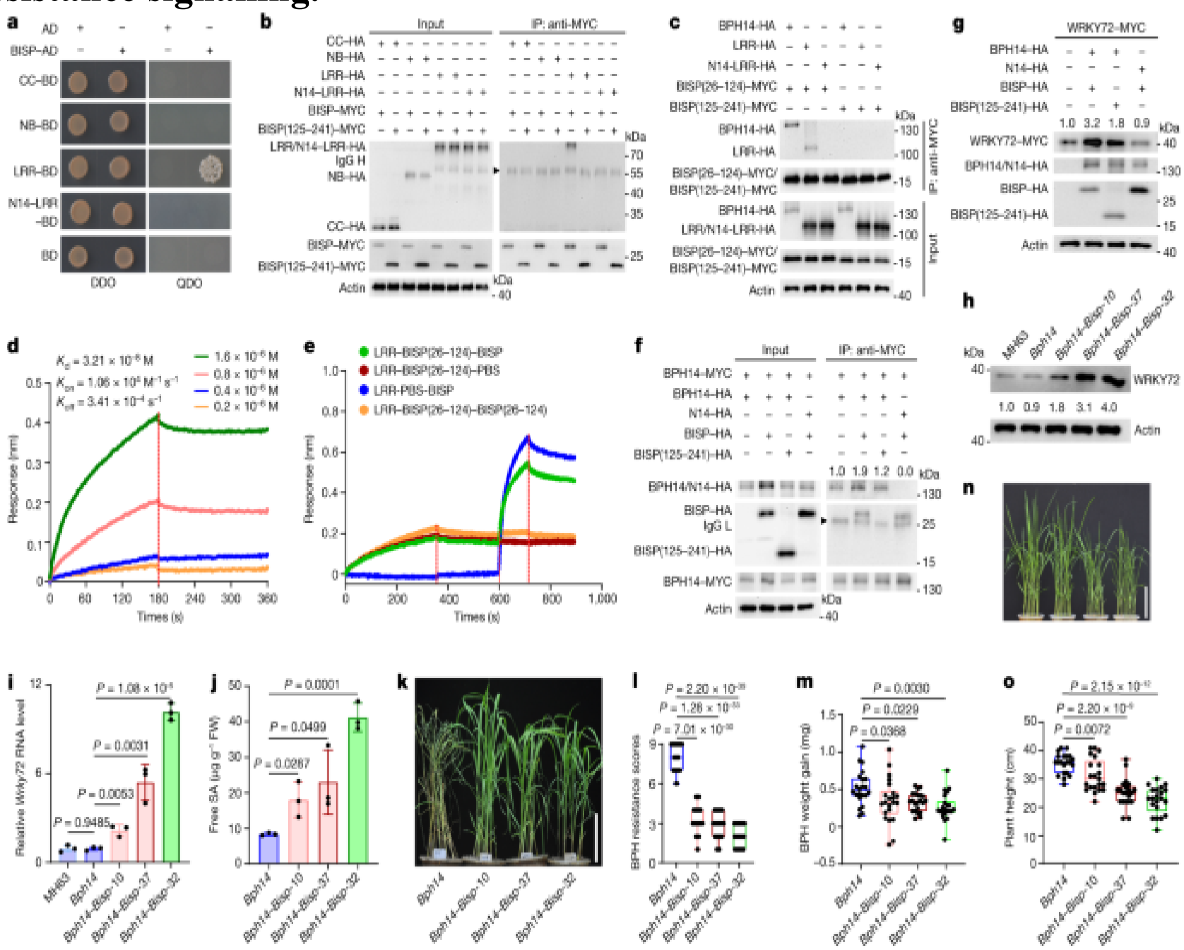
To further assess the function of BISP in BPH, we analysed the performance of BPHs treated with *Bisp*-RNAi or *GFP*-RNAi on Os*rlck185*-1 and WT plants. *Bisp*-RNAi insects showed significantly lower weight gain, honeydew excretion and survival rate than *GFP*-RNAi insects on Os*rlck185*-1 plants (Extended Data Fig. 2e–g). This result suggests that BISP has a specific function related to *Os*RLCK185 and may target other defence-related proteins in rice.

BISP activates BPH14 resistance

Pathogen effectors that are directly or indirectly recognized by NLR proteins trigger ETI[13,14]. Y2H and co-IP assays showed that BISP did not interact with either the CC or NB domain, but interacted with the LRR domain of BPH14 (Fig. 3a,b). Furthermore, the N terminus of BISP (amino acids 26–124) interacted with the LRR domain (Fig. 3c and Extended Data Fig. 5a).

Biolayer interferometry (BLI) assays using the BPH14–LRR or N14–LRR domains and BISP purified from insect cells showed that BPH14–LRR bound to BISP with a dissociation constant (K_d) of 3.21×10^{-8} M (Fig. 3d) and N14–LRR binding to BISP was not detected (Extended Data Fig. 5b). BPH14–LRR also bound to BISP(26–124) with a K_d of 1.27×10^{-7} M (Extended Data Fig. 5c), but not bind to BISP(125–241) (Extended Data Fig. 5d). Binding competition experiments showed that BISP competed with BISP(26–124) for binding to BPH14–LRR (Fig. 3e). Microscale thermophoresis (MST) assays confirmed that BISP directly bound to BPH14 (Extended Data Fig. 5e,f).

Fig. 3: BISP interacts with the BPH14 LRR domain and activates resistance signalling.



a,**b**, BISP interacted with the LRR domain of BPH14 in Y2H (**a**) and co-IP (**b**) assays. N14–LRR and BISP(125–241) served as negative controls. **c**, BISP(26–124) interacted with BPH14 and its LRR domain in a co-IP assay.

N14–LRR served as a negative control. **d**, BLI analysis for binding kinetics between LRR (BPH14–LRR) and BISP. BLI response profile for BISP at different concentrations with sensor-immobilized LRR. K_d, equilibrium dissociation constant; K_{on}, association rate constant; K_{off}, dissociation rate constant. **e**, BLI analysis for competitive binding between BISP and BISP(26–124) with LRR. **f**, BISP increased the levels of BPH14 homomeric complex in a co-IP assay in rice protoplasts. BISP(125–241) and N14 served as negative controls. Numbers above the lanes indicate band intensity relative to co-precipitated BPH14–MYC, quantified using ImageJ. **g**, Co-expression of BPH14 and BISP increased WRKY72 levels. **h**,**i**, Protein (**h**) and relative transcript (**i**) levels of *Wrky72* were increased in non-infested *Bph14–Bisp* lines compared with *Bph14* and MH63 plants. **j**, Endogenous free SA levels in non-infested *Bph14–Bisp* and *Bph14* plants. **k**,**l**, Phenotypes (**k**) and BPH-resistance scores (**l**) of *Bph14–Bisp* and *Bph14* plants after BPH infestation for 14 days. $n = 36$ plants examined over 3 independent experiments. **m**, Weight gain of BPHs feeding on *Bph14–Bisp* and *Bph14* plants for 48 h ($n = 22$, biologically independent samples). **n**,**o**, Photographs (**n**) and plant heights (**o**, $n = 21$ plants) of non-infested *Bph14–Bisp* and *Bph14* plants at the 4-leaf stage. In box plots in **l**, **m** and **o**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data and the centre line indicates the median. Individual data points are plotted. In **g** and **h**, numbers above/under the lanes indicate band intensity relative to actin (loading control) quantified using ImageJ. In **h**, **i** and **j**, data are the mean ± s.d. ($n = 3$, biologically independent experiments). In **i**, **j**, **l**, **m** and **o**, P values were derived by one-way ANOVA. Similar results were obtained from two (**d**,**e**) or three (**a**–**c**,**f–h**,**k**,**m**–**o**) independently replicated experiments. Scale bars, 10 cm (**k**,**n**).

Source data

BPH14 forms a homomeric complex and interacts with WRKY72 to increase WRKY72 accumulation, which in turn activates downstream defence signalling[19]. To examine the effects of BISP on the formation of BPH14 complexes and on WRKY72 protein levels, we co-expressed BPH14 and BISP in rice protoplasts. The levels of BPH14 self-association (Fig. 3f) and WRKY72 levels (Fig. 3g) were increased when BPH14 and BISP were

co-expressed in rice protoplasts, which suggested that BISP stimulated the activity of BPH14.

To examine the activation of BPH14 by BISP, we constructed transgenic *Bph14–Bisp* rice lines in which *Bisp* was constitutively expressed in the *Bph14* background (Extended Data Fig. 5g). WRKY72 protein and transcript levels were higher in three homozygous *Bph14–Bisp* plants relative to *Bph14* plants (Fig. 3h,i). In addition, non-infested *Bph14–Bisp* plants had higher levels of free SA than *Bph14* plants (Fig. 3j). Furthermore, the SA-responsive Os*ICS1* and Os*NPR1* and four rice *PR* genes were significantly upregulated (Extended Data Fig. 5h). The intensity of *Bph14*-mediated resistance responses in the *Bph14–Bisp* plants positively correlated with the levels of BISP protein.

We evaluated BPH resistance in *Bph14–Bisp* plants. Whereas *Bph14* plants showed little damage after 7 days of infestation (Fig. 1a), prolonged BPH feeding (14 days) caused substantial damage. By contrast, prolonged BPH feeding caused little damage to *Bph14–Bisp* plants (Fig. 3k,l), which suggested that ectopic expression of *Bisp* in *Bph14* plants enhances resistance to BPHs. In addition, BPH adults preferred *Bph14* over *Bph14–Bisp* plants in two-host choice assays and had poor performance on *Bph14–Bisp* plants based on BPH weight gain, honeydew excretion and survival (Fig. 3m and Extended Data Fig. 5i–k).

We further tested the performance of *Bisp*-RNAi and *GFP*-RNAi insects on *Bph14* plants and susceptible control MH63 plants (Extended Data Fig. 2h–j). The *GFP*-RNAi and non-injected insects exhibited significantly decreased weight gain, honeydew excretion and survival on *Bph14* plants compared with those on MH63 plants. By contrast, the *Bisp*-RNAi insects showed similar weight gain, honeydew excretion and survival rate when fed on *Bph14* or MH63 plants. Moreover, the *Bisp*-RNAi insects exhibited more weight gain, honeydew excretion and a higher survival rate than *GFP*-RNAi insects on *Bph14* plants. These results show that the knockdown of *Bisp* improved BPH performance on *Bph14* plants.

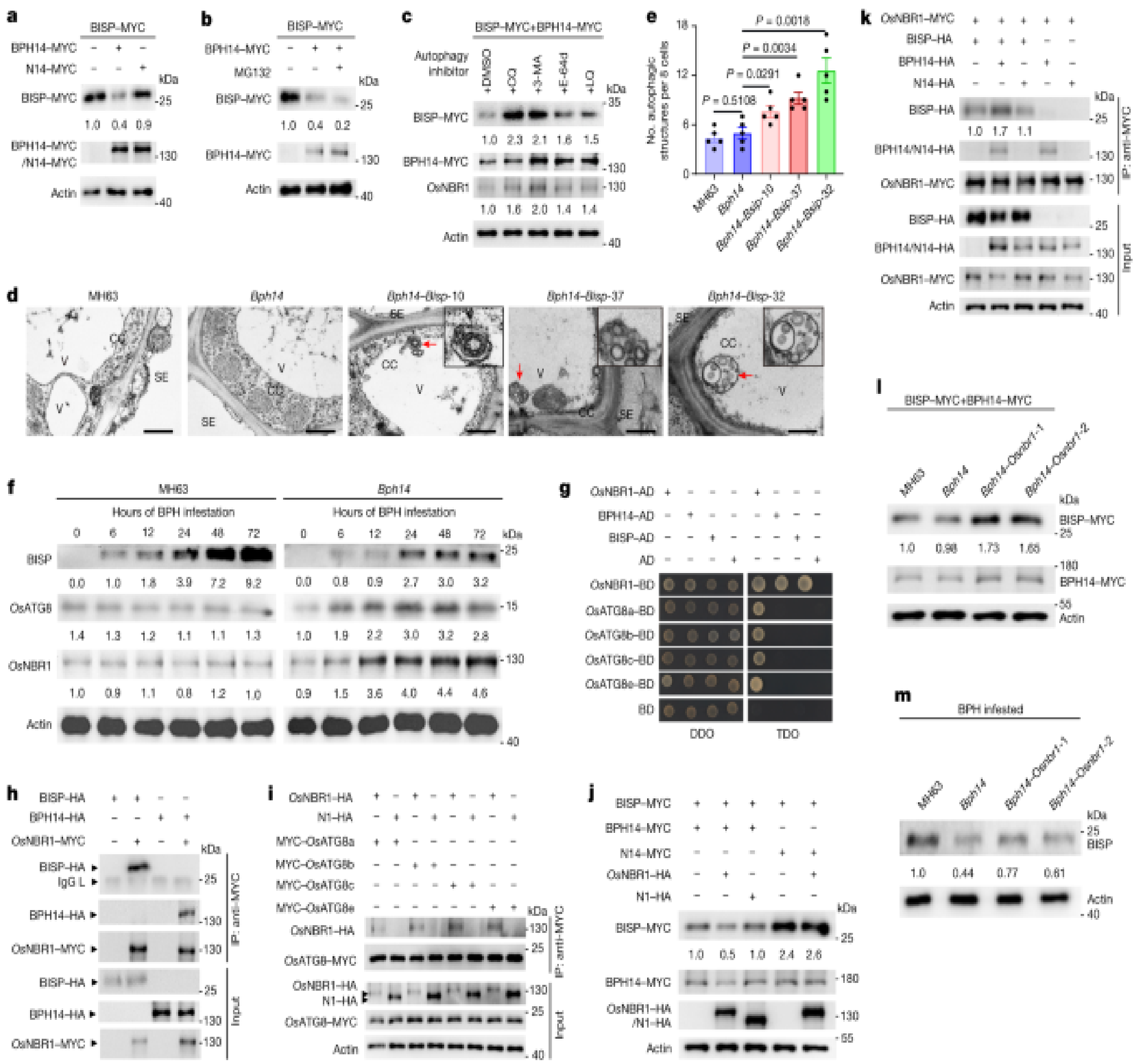
Constitutive activation of immune responses to pathogens usually negatively affects plant growth[23,24]. We observed that ectopic expression of *Bisp* in the *Bph14* background had significant fitness costs. *Bph14–Bisp* plants were

smaller than *Bph14* plants at both the seedling and heading stages (Fig. 3n,o and Extended Data Fig. 6a,b). Furthermore, the heading dates of *Bph14–Bisp* plants were advanced (Extended Data Fig. 6c). Finally, *Bph14–Bisp* plants exhibited poor agronomic traits, with significantly lower yields than *Bph14* plants (Extended Data Fig. 6d–j). The degree of fitness costs in *Bph14–Bisp* plants positively correlated with the levels of BISP and the intensity of *Bph14*-mediated resistance responses (Fig. 3h–m and Extended Data Figs. 5g–k and 6). These results indicate that constitutively expressed BISP induces BPH14-mediated resistance in *Bph14–Bisp* plants, which has a substantial fitness cost in rice.

BPH14 mediates BISP turnover by autophagy

The fitness costs imposed by ectopic expression of BISP in *Bph14* plants suggested that the activation of *Bph14*-mediated resistance should be tightly controlled in the natural habitats of rice. Several lines of evidence support this hypothesis. First, when BISP is expressed in rice protoplasts, BISP levels were lower when BISP was co-expressed with BPH14 (Fig. 4a), and BISP levels negatively correlated with BPH14 levels (Extended Data Fig. 7a). Indicative of the specificity of the BISP–BPH14 interaction, BISP levels were not reduced when BISP was co-expressed with N14 (Fig. 4a). These results indicate that BISP is degraded in a BPH14-dependent manner.

Fig. 4: BISP is degraded through *Os*NBR1-mediated autophagy.



a, Effects of BPH14 and N14 on BISP levels in rice protoplasts. **b**, Effects of the 26S proteasome inhibitor MG132 on BISP levels in rice protoplasts. **c**, Effects of autophagy inhibitors on BISP protein levels in rice protoplasts. 3-MA, 3-methyladenine; CQ, chloroquine; E-64d, aloxistatin; LQ, leupeptin. DMSO was the solvent for all inhibitors. **d**, Transmission electron microscopy images of autophagic structures in the phloem of non-infested *Bph14–Bisp* plants. Insets show enlarged autophagosome image at higher magnification. Red arrows indicate the location of double-membrane autophagosomes. CC, companion cell; SE, sieve element cell; V, vacuole. Scale bars, 500 nm. **e**, Quantification of double-membraned autophagosomes. *P* values were derived by one-way ANOVA. Data are the mean ± s.e.m. ($n = 5$, biologically independent experiments with every 8 cells as a biological replicate). **f**, Immunoblot detection of BISP, *Os*ATG8

and *Os*NBR1 in *Bph14* and MH63 plants. **g**, Y2H assay of the interactions between BISP, BPH14, *Os*NBR1 and four *Os*ATG8 proteins. **h**, Co-IP assays of interactions between *Os*NBR1 and BISP or BPH14 in rice protoplasts. **i**, Co-IP assay of interactions between *Os*NBR1 and four *Os*ATG8 proteins in rice protoplasts. The *Os*NBR1 mutant N1 served as a negative control. **j**, Effects of *Os*NBR1 and BPH14 on BISP levels in rice protoplasts. N14 and N1 served as negative controls. **k**, BPH14 enhanced interactions between *Os*NBR1 and BISP in rice protoplasts. N14 served as a negative control. Numbers under the lanes indicate band intensity relative to co-precipitated *Os*NBR1–MYC, quantified using ImageJ. **l**, Immunoblot detection of BISP in protoplasts of MH63, *Bph14* and *Os*NBR1 knockout (Os*nbr1*) plants. **m**, Immunoblot detection of BISP in MH63, *Bph14* and Os*nbr1* plants after BPH infestation. Numbers under the lanes (**a**–**c**,**f**,**j**,**l**,**m**) indicate protein abundance relative to that of actin (loading control), quantified using ImageJ. Experiments (**a**–**d**,**f**–**m**) were repeated three times and gave similar results.

Source data

The ubiquitin–proteasome system and autophagy are the main protein-degradation pathways in plant cells[25]. The ubiquitin–proteasome system is a key post-translational mechanism for controlling plant immune responses to pathogens[26]. To test whether the ubiquitin–proteasome system is involved in the modulation of BISP–BPH14 immune activation, the 26S proteasome inhibitor MG132 was added to protoplasts expressing BISP and BPH14. MG132 blocked the proteasome pathway degradation of WRKY72 (Extended Data Fig. 7b), but did not inhibit the degradation of BISP (Fig. 4b). By contrast, treatment with four autophagy inhibitors blocked the degradation of BISP (Fig. 4c and Extended Data Fig. 7c), which suggested that the autophagy pathway is involved in the degradation of BISP.

To further assess the role of autophagy in BISP turnover, we used cyan fluorescent protein-tagged ATG8F (CFP–ATG8F) as a marker[27,28] to monitor autophagy (Extended Data Fig. 7d,e). When CFP–ATG8F was expressed with BISP or BPH14 individually in *Nicotiana benthamiana* leaves (hereafter *Nb* is used to denote *N. benthamiana*-related proteins or genes), few punctate CFP structures, corresponding to pre-autophagosomes and autophagosomes, were detected. However, when CFP–ATG8F was

expressed with BISP and BPH14, the number of punctate CFP structures significantly increased, which indicated that BISP and BPH14 co-expression induced autophagy. This interaction was specific, as expressing CFP–ATG8F with BISP and N14 did not increase the number of autophagosomes. Concanamycin A (ConA) blocks autophagic flux[28], and the numbers of autophagic bodies were substantially increased in cells that expressed both BISP and BPH14 and treated with ConA. When we silenced the core autophagy genes Nb*ATG6*, Nb*PI3K* and Nb*ATG7*, the numbers of autophagic bodies were markedly decreased in the silenced leaves that expressed CFP–ATG8F with BISP–yellow fluorescent protein (YFP) and BPH14 (Extended Data Fig. 7f–h). Furthermore, using transmission electron microscopy, we observed more double-membraned autophagosome structures in the phloem companion cells of non-infested *Bph14–Bisp* plants than in *Bph14* and MH63 plants (Fig. 4d,e). Accordingly, *Os*ATG8 protein levels and Os*ATG8a*, Os*ATG8b* and Os*ATG8c* transcript levels were significantly higher in non-infested *Bph14–Bisp* plants than in *Bph14* and MH63 plants (Extended Data Fig. 7i,j). These observations indicate that BISP is degraded through autophagy in a BPH14-dependent manner.

As BPHs secrete saliva containing BISP into rice tissues when feeding (Fig. 1j,k and Extended Data Fig. 1q–t), we monitored BISP levels in infested BPH-susceptible (MH63) or BPH-resistant (*Bph14*) plants. In MH63 plants, BISP was first detected at 6 h and increased continuously for 72 h of BPH feeding (Fig. 4f). By contrast, BISP levels were lower in *Bph14* plants, with levels sustained after 24 h of feeding (Fig. 4f). *Os*ATG8 levels did not significantly change in BPH-infested susceptible plants. By contrast, in *Bph14* plants, *Os*ATG8 levels increased by 6 h after feeding and were maintained for 72 h (Fig. 4f). These results suggest that autophagy is activated in *Bph14* plants but not in BPH-susceptible rice. These data correlated well with the upregulation of three Os*ATG8* genes in *Bph14* plants compared with MH63 plants following BPH infestation (Extended Data Fig. 7k). ConA treatment substantially increased levels of ATG8 proteins in BPH-infested *Bph14* plants (Extended Data Fig. 7l). These results indicate that BPH feeding activates autophagy and that BISP levels are tightly controlled in *Bph14* plants.

*Os*NBR1-mediated autophagy controls BISP levels

Mechanisms of autophagy are conserved in yeast, plants and metazoans[29]. In rice and other plants, multiple ATG8 proteins mediate autophagy[30,31,32]. To clarify the molecular mechanisms that underlie the BPH14-dependent degradation of BISP, we examined the interactions of four *Os*ATG8 proteins with BISP or BPH14 using Y2H assays. Neither BISP nor BPH14 interacted with the *Os*ATG8 proteins (Fig. 4g and Extended Data Fig. 8a).

In addition to canonical autophagy, selective autophagy targets specific cargoes, such as intracellular pathogens and defective proteins, and plays pivotal roles in eukaryotic organisms[33,34,35]. NBR1 is a selective autophagy cargo receptor and interacts with ATG8 proteins to facilitate the autophagy-dependent degradation of target proteins[33,34,35,36,37]. We therefore examined the interaction of BISP or BPH14 with *Os*NBR1 in Y2H assays. Both BISP and BPH14 interacted with *Os*NBR1 (Fig. 4g and Extended Data Fig. 8a). Co-IP assays in rice protoplasts confirmed these interactions (Fig. 4h). Furthermore, *Os*NBR1 interacted with all four *Os*ATG8 proteins in Y2H assays (Fig. 4g and Extended Data Fig. 8a) and in rice protoplasts (Fig. 4i), which suggested that *Os*NBR1 might mediate the degradation of BISP.

The roles of *Os*NBR1 in the autophagic degradation of BISP were examined by immunoblotting. The expression of BISP with *Os*NBR1 in rice protoplasts did not influence BISP levels. By contrast, the expression of BISP, *Os*NBR1 and BPH14 reduced the level of BISP (Fig. 4j), which suggested that *Os*NBR1 facilitated BPH14-dependent degradation of BISP. We used CFP–ATG8F to determine the effects of *Os*NBR1 on the numbers of autophagic vesicles. The expression of *Os*NBR1 with BISP and BPH14 induced more punctuate CFP–ATG8F structures in *N. benthamiana* leaves than expression of only BISP and BPH14 (Extended Data Fig. 8b,c). Additionally, *Os*NBR1 protein and transcript levels were significantly higher in non-infested *Bph14–Bisp* plants than in *Bph14* and MH63 plants (Extended Data Fig. 8d,e). Following BPH infestation, *Os*NBR1 protein and RNA levels were increased in *Bph14* plants, but not in BPH-susceptible MH63 plants (Fig. 4f and Extended Data Fig. 8f). Moreover, ConA treatment significantly increased *Os*NBR1 levels in *Bph14* plants following BPH infestation (Extended Data Fig. 7l). Y2H and co-IP assays showed that the CC and LRR domains of BPH14 interacted with *Os*NBR1 (Extended

Data Fig. 9a,b). Co-IP assays also showed that BPH14 enhanced the interaction of *Os*NBR1 with BISP (Fig. 4k).

NBR1 homopolymerizes through its PB1 domain, binds to ubiquitin through its UBA domain and interacts with ATG8 through a conserved LIR motif to promote selective autophagy[37]. *Os*NBR1 contains these interaction domains, as well as conserved ZZ-type zinc finger and BRCA1 domains. To determine which *Os*NBR1 domain (or domains) participates in the degradation of BISP, we generated four truncated forms (N1 to N4) and three site-directed mutants (N5 to N7) of *Os*NBR1 (Extended Data Fig. 9c). Co-expression assays showed that none of mutants were able to degrade BISP (Extended Data Fig. 9d). Whereas the Os*nbr1* N5 mutant lost the ability to form homopolymers (Extended Data Fig. 9h), the other Os*nbr1* mutants retained the ability to form homopolymers. However, these mutants failed to interact with BPH14, BISP or *Os*ATG8C (Extended Data Fig. 9e–i). These results suggest that *Os*NBR1 is involved in BISP degradation through autophagy and that domains outside the UBA in *Os*NBR1 are important for interacting with its target BISP.

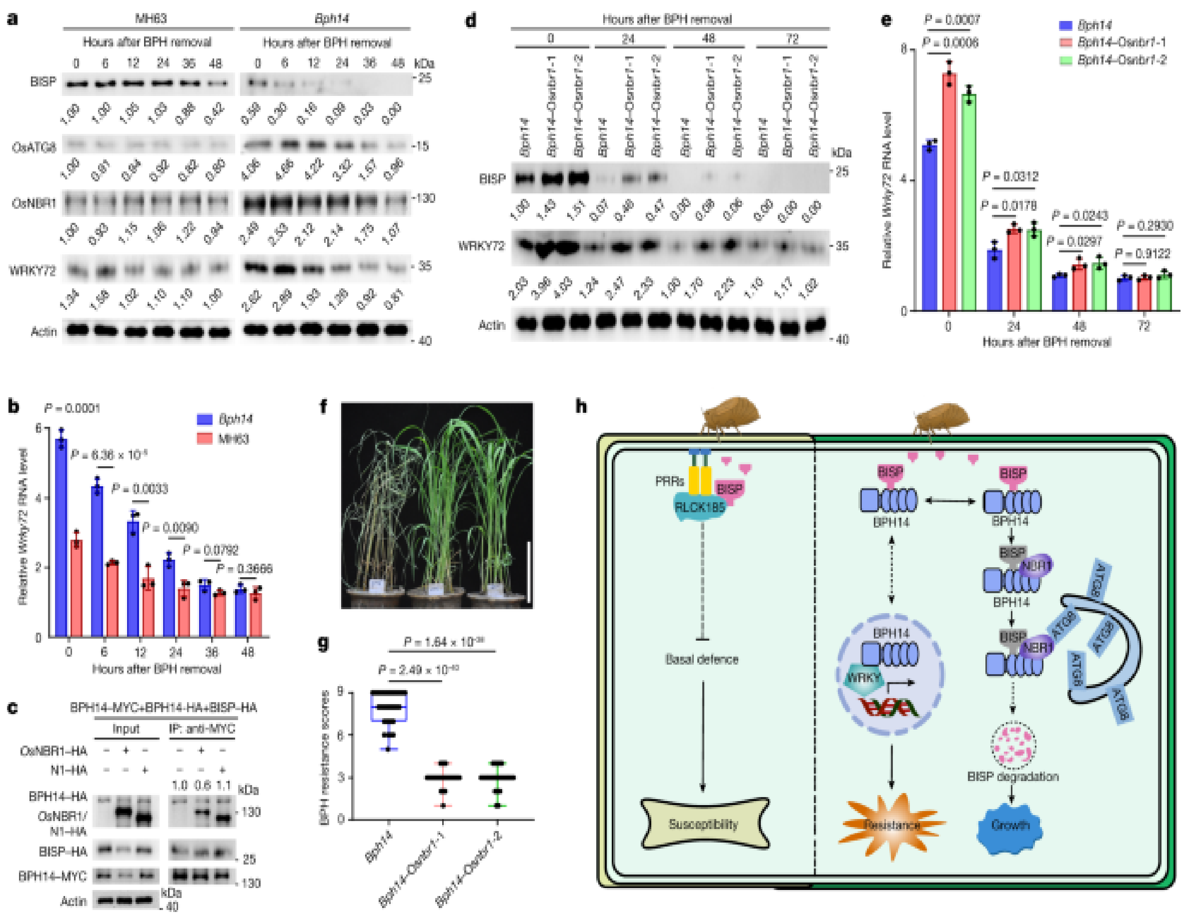
We generated *Os*NBR1 knockout mutants in the *Bph14* background (*Bph14*–Os*nbr1*) to further analyse the roles of *Os*NBR1 in the degradation of BISP (Extended Data Fig. 9j). *Os*NBR1 was undetectable in the *Bph14*–Os*nbr1* mutants (Extended Data Fig. 9k). We then expressed BISP and BPH14 in *Bph14*–Os*nbr1*, *Bph14* and MH63 protoplasts. Consistent with a role for *Os*NBR1 in BISP turnover, BISP levels were higher in *Bph14*–Os*nbr1* protoplasts than in *Bph14* and MH63 protoplasts (Fig. 4l). BISP was more abundant in infested *Bph14*–Os*nbr1* plants than in *Bph14* plants, but less abundant than in MH63 plants (Fig. 4m). These results indicate that *Os*NBR1-mediated autophagic degradation controls BISP levels.

BISP degradation attenuates resistance

On susceptible plants, BPHs consume phloem sap for prolonged times[17]. By contrast, on BPH-resistant plants, various resistance factors prevent sustained phloem ingestion[11,12,16,17,38]. BPHs have a higher frequency of stylet probing, penetrating numerous sites in an attempt to find an acceptable feeding site[17,39].

To provide insights into the duration of *Bph14*-mediated immunity, we assessed the longevity of BISP and the immune status of rice plants after BPH feeding had ceased. To this end, BPH-resistant *Bph14* plants and BPH-susceptible MH63 plants were infested with BPHs for 24 h. At this time, insects were removed, and BISP, *Os*ATG8, *Os*NBR1 and WRKY72 levels were monitored for 48 h. In MH63 plants, BISP levels were similar at 0 and 36 h after BPH removal. By contrast, in *Bph14* plants, BISP levels decreased markedly at 6 h and were barely detected at 36 h (Fig. 5a). Correlating well with the decreased BISP levels, the levels of *Os*ATG8 and *Os*NBR1 proteins and their transcripts were more abundant in *Bph14* plants than in MH63 plants (Fig. 5a and Extended Data Fig. 10a). These results suggest that after BPH feeding and secretion of saliva have ceased, BISP is stable for prolonged times in susceptible rice but is rapidly eliminated through autophagy in *Bph14* plants.

Fig. 5: BPH14-mediated BISP degradation contributes to the termination of BPH14 activation.



a, Immunoblot detection of BISP, *Os*ATG8, *Os*NBR1, and WRKY72 in *Bph14* and MH63 plants after cessation of BPH feeding. **b**, Relative *Wrky72* transcript levels in *Bph14* and MH63 plants after cessation of BPH feeding. **c**, Co-IP detection of levels of BPH14 homomeric complex formation. N1 served as a negative control. Numbers above the lanes indicate band intensity relative to co-precipitated BPH14–MYC, quantified using ImageJ. **d**, Immunoblot detection of BISP and WRKY72 in *Bph14* and *Bph14*–Os*nbr1* plants after cessation of BPH feeding. **e**, Relative *Wrky72* transcript levels in *Bph14* and *Bph14*–Os*nbr1* plants after cessation of BPH feeding. **f**,**g**, Photographs (**f**) and resistance scores (**g**) of *Bph14* and *Bph14*–Os*nbr1* plants after 14 days of BPH infestation. Scale bar, 10 cm. The box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data and the centre line indicates the median. Individual data points are plotted ($n = 42$ plants examined over 3 independent experiments). **h**, A working model for BISP-induced immunity regulation. In BPH-susceptible rice, BISP interacts with *Os*RLCK185 and suppresses its phosphorylation activity, thereby inhibiting basal defence and promoting BPH feeding. In rice carrying the BPH resistance gene *Bph14*, BISP binds directly to the LRR domain of BPH14 and activates HPR. Continuous activation of HPR is detrimental to plant growth and reproduction. Therefore, following activation of BPH14-mediated immunity, BISP–BPH14 binds to *Os*NBR1, which mediates the autophagic degradation of BISP. The degradation of BISP leads to the termination of BPH14-mediated immunity, which restores cellular homeostasis and prevents the fitness costs associated with prolonged activation of BPH14-mediated immunity. The timing of dissociation of the BISP–BPH14 complex after association with *Os*NBR1 and the fate of BPH14 are yet to be determined. Numbers under lanes (**a**,**d**) indicate protein abundance relative to that of actin (loading control, **a**,**c**,**d**), quantified using ImageJ. In **b**, **c**, **e** and **g**, *P* values were derived by one-way ANOVA. In **b** and **e**, data are the mean ± s.d. ($n = 3$, biologically independent experiments). Experiments (**a**,**c**,**d**,**f**) were repeated three times, each giving similar results. PRRs, pattern recognition receptors.

Source data

Moreover, the degradation of BISP was accompanied by a mitigation of the signal outputs of BPH14-activated resistance (Fig. 5a,b and Extended Data Fig. 10b). For example, WRKY72 protein and transcript levels rapidly

reduced (within 12 h) in *Bph14* plants and were slower in susceptible plants after BPH feeding had ceased (Fig. 5a,b). In addition, Os*ICS1* and Os*NPR1* transcripts showed a similar regulation pattern (Extended Data Fig. 10b). The rapid dissipation of BPH14-regulated immunity based on BISP depletion was supported by the dynamics of formation of the BPH14 homomeric complex, which was reduced when BPH14 was expressed with BISP and *Os*NBR1 to induce autophagy (Fig. 5c). These results suggest that autophagy contributes to the elimination of BISP and the attenuation of resistance in *Bph14* plants after BPH feeding is terminated.

Further support for this theory was garnered from examining the fate of BISP and defence-signalling components in *Bph14*–Os*nbr1* plants and WT *Bph14* plants after BPH feeding was terminated. BISP levels were significantly higher in *Bph14*–Os*nbr1* plants than in *Bph14* plants at 0, 24 and 48 h after removing the insects (Fig. 5d), which confirmed that *Os*NBR1 deficiency slowed the degradation of BISP. WRKY72 protein and transcript levels, as well as Os*ics*1 and Os*npr*1 transcripts, were significantly higher at 0–48 h in the *Bph14*–Os*nbr1* plants than the *Bph14* plants (Fig. 5d,e and Extended Data Fig. 10c). This result suggests that the time span of BPH14-activated signalling is extended in *Bph14*–Os*nbr1* plants.

Finally, we examined the impact of *Os*NBR1 on rice resistance to BPHs. Whereas *Bph14* plants were damaged by prolonged BPH infestation (14 days), *Bph14*–Os*nbr1* plants had limited symptoms and displayed a higher level of BPH resistance (Fig. 5f,g). BPH adults preferred *Bph14* plants to *Bph14*–Os*nbr1* plants in a two-host choice test. On the basis of honeydew excretion, weight gain and survival, BPHs had poorer performance when feeding on *Bph14*–Os*nbr1* plants (Extended Data Fig. 10d–g). These results indicate that *Os*NBR1 deficiency enhances BPH resistance in *Bph14* plants.

Discussion

HPR conferred by NLR proteins is widely used for controlling insect pests[40]. Here we provide molecular and cellular insights into the long-standing question of how plant NLR receptors perceive insect-feeding signals to deploy and modulate ETI resistance mediated by these proteins. Notably, we discovered that the tripartite interactions of BISP, its cognate

NLR receptor BPH14 and the selective-autophagy receptor NBR1 enable a self-regulated resistance mechanism to BPHs (Fig. 5h).

In this tripartite interaction system, the BISP effector is secreted into rice leaf sheaths to bind to one of its targets, *Os*RLCK185. By interfering with *Os*RLCK185 autophosphorylation, BISP suppresses rice immune responses to promote BPH feeding and success. However, in resistant rice, the NLR receptor BPH14 directly binds to BISP. Following recognition, resistance is activated, which stops BPHs from ingesting phloem sap[8,17,19]. Constitutive activation of BISP-triggered resistance in *Bph14* plants is detrimental to plant fitness, reducing plant stature and yields. Therefore, mechanisms to temper the resistance response must be deployed to promote plant vitality. The importance of the ubiquitin–proteasome system in effector and NLR protein turnover in pathogen–plant ETI is known[26,41]. Here we revealed that the governor of *Bph14*-mediated resistance is *Os*NBR1-regulated autophagy, which controls BISP turnover. Both BISP and BPH14 physically interact with *Os*NBR1. In turn, *Os*NBR1 interacts with *Os*ATG8 to promote the degradation of BISP and, thereby, fine-tune the level of *Bph14*-mediated resistance. Whether the BISP–BPH14 complex dissociates before BISP turnover by autophagy or whether BPH14 is also degraded by autophagy is an essential topic of research in the future. Finally, although autophagy can be induced by transient reprogramming in response to plant hormones[42], the role of SA, which is induced in the BISP–BPH14 interaction, remains to be elucidated.

The BISP–BPH14–*Os*NBR1 resistance response adapts to changes in BPH feeding pressure. BPHs terminate feeding when a host becomes undesirable, which can induce individual insects or entire BPH populations to migrate long distances in search of better host plants[15,16,17,18]. When BPHs feed on *Bph14* plants, BISP triggers resistance, which makes *Bph14* plants unpalatable and forces the insects to stop feeding. With the cessation of feeding and the departure of insects from these sites, the activated autophagy pathway rapidly degrades BISP to restore cellular homeostasis. Immune systems resume their ‘off’ status, thereby allowing rice to reallocate resources to growth and reproduction. The newly described BISP–BPH14–*Os*NBR1 tripartite interaction system enables rice to resist BPHs without compromising yield performance in the natural habitats[43]. By leveraging the

autoregulatory mechanisms used in *Bph14* gene-for-gene resistance in new ways, we have the ability to create a new generation of resistant rice that can feed the increasing world population in a sustainable manner and, significantly, to control insect pests without relying on insecticide sprays.

Methods

Plant materials and growth conditions

Nipponbare (designated as *N14*) is a model *O. sativa japonica* rice variety that contains the *N14*-susceptible allele of *Bph14*. *N14* was used in early experiments to establish the principles of BISP–BPH14 interactions. The RI35 line (designated as *Bph14*) is a recombinant inbred line that contains the BPH resistance gene *Bph14* (ref. [8]). Minghui 63 (MH63) is the BPH-susceptible parent of RI35 (ref. [44]) and is a model variety for *O. sativa indica* rice breeding and genomics. MH63 and BPH14 were used for most studies to elucidate BISP–BPH14–NBR1 interactions owing to the agronomic importance of MH63. The following transgenic lines were developed in this study: *N14–Bisp* (*pUBI::Bisp-Myc*, *N14* recipient), *Bph14–Bisp* (*pUBI::Bisp–Myc*, *Bph14* recipient), Os*rlck185* (*Os*RLCK185–Cas9, ZH11 recipient), *Bph14*–Os*nbr1* (*Os*NBR1–Cas9, *Bph14* recipient). The Os*rlck185* mutants were generated in the ZH11 background[45]; ZH11 is a readily transformable *O. sativa japonica* rice that BPH-susceptible genotype.

Insect materials and growth conditions

The BPH insects were maintained on susceptible cultivar Taichung Native 1 at 28 ± 2 °C with 60–80% relative humidity and a photoperiod of 16 h of light–8 h of dark. RNAi of BPH by dsRNA injection was developed as part of this study: *GFP*-RNAi (microinjected with ds*GFP*) and *Bisp*-RNAi (microinjected with ds*Bisp*).

BPH bioassays on rice plants

To evaluate BPH resistance of transgenic and WT rice plants, sets of around 15 or 20 seeds from the same plant were sown in a plastic cup (10 cm in

diameter, 20 cm in height). At the three-leaf stage, all plants were placed under a large gauze cover, and each seedling was infested with ten second- or third-instar BPH nymphs. When all of the susceptible plants had died (scored as 9), each seedling of the other cultivars or lines was given a score of 0, 1, 3, 5, 7, or 9 according to the degree of damage, as previously described[8,46,47], and the BPH resistance score was calculated. At least three replicates were used for each cultivar or line.

To measure the weight gain and honeydew excretion of BPHs, a newly emerged short-winged female adult and a Parafilm sachet were weighed using a 1/100,000 Shimadzu AUW120D electronic balance[11,12]. The insect was placed in the pre-weighed sachet, which was then attached to the leaf sheath near the lower part of a rice plant. After 48 h, the insect and Parafilm sachet were weighed again. The difference between the first and second measurements of the weight of the insect was recorded as the BPH weight gain; the difference between the first and second measurements of the weight of the sachet was recorded as amount of honeydew excreted. At least ten replicates were used for each cultivar or line.

For the two-host choice tests, a transgenic plant and a WT plant were grown in a plastic cup as described elsewhere[10,11,12]. At the four-leaf stage, the cup was covered with gauze, and 20 or 30 second- or third-instar nymphs were released into the cup (the exact number depended on the BPH resistance of the WT plant). The number of nymphs that settled on each plant was recorded at 6, 12, 24, 48, 72, 96, and 120 h after release (the moment of release was considered time 0). Ten cups were analysed for each cultivar or line examined.

To study BPH nymph survival on rice, second- or third-stage BPH nymphs were released (ten insects per plant), and the cups were covered with a light-transmitting mesh. The nymphs on each plant were counted 8 or 9 days after release. Survival rates were calculated as the number of surviving nymphs divided by the total number of nymphs released at the start of the experiment. Ten cups were analysed for each cultivar or line.

Plasmid construction

Routine molecular cloning techniques were used to prepare the constructs. The plasmids and primers used in this work are listed in Supplementary Table 3. All of the resulting recombinant vectors were sequenced.

To construct the plasmid for Y2H screening, the full-length *Bph14* sequence was amplified from *Bph14* leaf sheath cDNA and cloned into the *Xma* I site of the pGBKT7 vector, which produced the BPH14–BD vector.

To construct the plasmids for the Y2H assays, the coding sequences of *Bisp* (ORF sequence without its 1–25 amino acid residue signal peptide; amino acids 26–241), *Bisp*$^{26–124}$(amino acids 26–124) and *Bisp*$^{125–241}$ (amino acids 125–241) were amplified and cloned into the *Xma* I site of the pGADT7 vector to produce BISP–AD, BISP(26–124)–AD and BISP(125–241)–AD, respectively. The coding sequences of the CC, NB and LRR domains and full-length *Bph14* were amplified from the BPH14-BD vector and cloned into the *Xma* I sites of the pGADT7 and pGBKT7 vector to produce CC–AD, NB–AD, LRR–AD, BPH14–AD, CC–BD, NB–BD, and LRR–BD, respectively. The coding sequence of *N14* (*Bph14* allele in the rice variety Nipponbare) was amplified from *N14* leaf sheath cDNA and cloned into the *Xma* I site of the pGBKT7 vector to produce N14–BD.

Eight kinases that were differentially phosphorylated during BPH-infested and non-infested plants were chosen for this study. The coding sequences for the Os*RLCK185*, Os*RLCK185*$^{1–85}$ (amino acids 1–85), Os*RLCK185*^{KD} (kinase domain, amino acids 86–353), Os*RLCK185*$^{354–491}$ (amino acids 354–491), Os*MAPKKKε*, Os*MKK6*, Os*MPK20*, Os*MPK16*, Os*CDPK14*, Os*CDPK20*, and Os*RLCK259* fragments were amplified from *N14* leaf sheath cDNA and cloned into the pGBKT7 vector to produce *Os*RLCK185–BD, *Os*RLCK185(1–85)–BD, *Os*RLCK185^{KD}–BD, *Os*RLCK185(354–491)–BD, *Os*MAPKKKε–BD, *Os*MKK6–BD, *Os*MPK20–BD, *Os*MPK16–BD, *Os*CDPK14–BD, *Os*CDPK20–BD, and *Os*RLCK259–BD, respectively. The coding sequences of Os*NBR1*, Os*ATG8a*, Os*ATG8b*, Os*ATG8c*, and Os*ATG8e* were amplified from *Bph14* leaf sheath cDNA and cloned into pGBKT7 and pGADT7 to produce *Os*NBR1–BD, *Os*ATG8A–BD, *Os*ATG8B–BD, *Os*ATG8C–BD, and *Os*ATG8E–BD, respectively.

Constructs used for rice protoplast transfection were generated using pCXUN-4×HA and pCXUN-4×Myc[19]. The coding sequences of the CC, NB

and LRR domains, *Bph14*, *N14*, the N14 LRR domain (N14–LRR), *Wrky72, Bisp* (ORF sequence without its signal peptide), *Bisp*$^{26–124}$, *Bisp*$^{125–241}$, Os*RLCK185*, Os*RLCK259*, Os*NBR1*, Os*NBR1* mutant N1, Os*ATG8a*, Os*ATG8b*, Os*ATG8c*, and Os*ATG8e* were amplified and cloned into the pCXUN-4×Myc vector to produce CC–MYC, NB–MYC, LRR–MYC, BPH14–MYC, N14–MYC, N14–LRR–MYC, WRKY72–MYC, BISP–MYC, BISP(26–124)–MYC, BISP(125–241)–MYC, *Os*RLCK185–MYC, *Os*RLCK259–MYC, N1–MYC, *Os*NBR1–MYC, MYC–*Os*ATG8A, MYC–*Os*ATG8B, MYC–*Os*ATG8C, and MYC–*Os*ATG8E, respectively. Meanwhile, the fragments for the GFP, CC, NB, LRR domains, *Bph14*, *N14*, the N14 LRR domain (N14–LRR), *Bisp*, *Bisp*$^{26–124}$, *Bisp*$^{125–241}$, Os*RLCK185*, and Os*NBR1* were amplified and cloned into the pCXUN-4×HA vector to produce GFP–HA, CC–HA, NB–HA, LRR–HA, BPH14–HA, N14–HA, N14–LRR–HA, BISP–HA, BISP(26–124)–HA, BISP(125–241)–HA, *Os*RLCK185–HA, and *Os*NBR1–HA, respectively. Additionally, we generated a series of Os*nbr1* fragments, including the isolated N1 (amino acids 1–755), N2 (amino acids 756–832), N3 (amino acids 1–782), N4 (amino acids 1–755 and 798–832), and mutant N5 (K13A), N6 (WL788/791AA) and N7 (K13A and WL788/791AA) fragments and cloned them into the pCXUN-4×HA vector to produce N1–HA, N2–HA, N3–HA, N4–HA, N5–HA, N6–HA, and N7–HA, respectively.

To analyse the subcellular localization of BISP, the coding sequence of *Bisp* (ORF sequence without its signal peptide) was cloned downstream of the maize (*Zea mays*) *Ubiquitin 1* promoter, in-frame with *GFP* in the binary vector pCAMBIA1300, which produced the BISP–GFP construct.

For BiFC analysis, the coding sequences of *Bph14*, *N14*, *Bisp* (ORF sequence without its signal peptide), and *Bisp*$^{125–241}$ were cloned in-frame into the *Xma* I sites of the pUSYNE and pUSYCE vectors[19], which produced the constructs BISP–YN, BISP(125–241)–YN, BPH14–YC, and N14–YC, respectively.

The *Escherichia coli* recombinant protein vectors used for the expression and purification of proteins in the phosphorylation activity assays, designated pET-MBP-His and pET-GST-His, were created by adding a C-terminal MBP or GST tag, respectively, to the pET-28a expression vector

(EMD Biosciences, Novagen). The coding sequences of *Bisp* (ORF sequence without its signal peptide), *Bisp*$^{26–124}$ and Os*RLCK185* were amplified and cloned into the pET-MBP-His and pET-GST-His vector, which produced *Os*RLCK185–MBP–His, BISP(26–124)–GST–His and BISP–GST–His constructs, respectively.

For the *N. benthamiana* leaf agroinfiltration experiments, the coding sequences of *Bph14*, *N14*, Os*NBR1,* and N1 (an *Os*NBR1 mutant) were cloned into the pEarleyGate 203 vector[48] to produce BPH14-203, N14-203, *Os*NBR1-203, and N1-203, respectively. The coding sequence of *Bisp* (ORF sequence without its signal peptide) was cloned into the pEarleyGate 201 and pEarleyGate 101 vector[48] to produce BISP-201 and BISP-101 (BISP–YFP), respectively.

To generate the CRISPR–Cas9 construct, the target sites for CRISPR–Cas9 were designed using the web-based tool CRISPR-P (v.2.0; http://crispr.hzau.edu.cn/CRISPR2/). The CRISPR–Cas9 binary construct was prepared as previously described[49], which produced the construct *Os*NBR1–Cas9.

Y2H screening

The Matchmaker Gold Yeast Two-Hybrid System (Clontech, 630489) was used to screen for BPH14-interacting proteins. Y2H Gold cells carrying BPH14–BD were mixed with 1 ml of a BPH salivary gland cDNA library and incubated overnight before plating on TDO (SD/-Leu-Trp-His)-selective medium. Candidate clones growing on TDO medium were confirmed following the manufacturer's protocol (Clontech, 630489). The BPH cDNA library was constructed using the Make Your Own "Mate & Plate" Library System (Clontech, 630490) with a simple and highly efficient protocol. Total RNA was isolated from BPH salivary glands using TRIzol reagent (Takara, 9109), and cDNA was synthesized from the isolated mRNA using SMART cDNA synthesis technology (Clontech, 634926).

For the Y2H assay of protein interactions, the yeast strain AH109 was transformed with the indicated bait and prey plasmids according to the manufacturer's instructions. Co-transformants were simultaneously plated

on selection medium containing DDO (SD/-Leu-Trp), TDO (SD/-Leu-Trp-His) with the appropriate concentration of 3-amino-1,2,4-triazole (3-AT) or QDO (SD/-Leu-Trp-His-Ade) and incubated at 30 °C.

Cloning and sequencing of *Bisp* cDNA

The cDNA sequences of *Bisp* were obtained from BPH salivary gland transcriptomes[5]. To obtain the full-length counterparts of the truncated sequences in the transcriptomes, 5′- and 3′-RACE amplification was performed using a SMARTer RACE cDNA Amplification kit (Clontech, 634923) following the manufacturer's instructions. *Bisp*-specific primers were obtained based on sequencing data. The RACE products were amplified, and the purified products were ligated into the pMD18-T vector (Takara, 6011) and sequenced.

Protein extraction, immunoblot analysis and in vivo co-IP assays

Proteins were extracted from the leaf sheaths of rice seedlings at the four-leaf stage. The leaf sheaths were wiped with a cotton ball soaked in alcohol to remove honeydew and ground in liquid nitrogen. Total protein was extracted from the leaf tissue in rice protein extraction buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol, 0.1% NP-40, plant protease inhibitor cocktail (Roche)). Equal amounts of total protein were analysed by SDS–PAGE and detected by immunoblotting using anti-actin (Abbkine, ABL1050; 1:3,000), anti-BISP (1:500), anti-*Os*WRKY72 (Beijing Protein Innovation, AbP80456-A-SE; 1:500) and anti-*At*NBR1 (Agrisera, AS194281; 1:1,000) antibodies. For *Os*ATG8 analysis, equal amounts of proteins were subjected to SDS–PAGE with 6 M urea[50] and detected by immunoblotting using anti-*At*ATG8A antibodies (Abcam, ab77003; 1:1,000). The anti-BISP and anti-NISP1 antibodies were prepared by expressing BISP and NISP1 (cloned into pET28a) in *E. coli* strain BL21 (DE3), respectively. The expressed recombinant proteins were collected and injected into two rabbits (DIA·AA Biotech). Ponceau S solution (Sigma-Aldrich, P7170) was used to stain the PVDF membrane and was used as a loading control. The intensities of protein signals were quantified using ImageJ.

Protein samples from rice protoplasts were prepared in rice protein extraction buffer (100 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 mM $MgCl_2$, 0.5% (w/v) Triton X-100, with a plant protease inhibitor cocktail (Roche)). Total soluble proteins were extracted from rice protoplast samples, each in 100 µl of rice protoplast protein extraction buffer. Next, 10 µl of the extract was separated by SDS–PAGE and subjected to immunoblotting using anti-HA (MBL, M180-3; 1:1,000), anti-MYC (MBL, M192-3; 1:1,000) or anti-actin (Abbkine, ABL1050; 1:3,000) antibodies.

For the co-IP assays, rice protoplasts were incubated at 28 °C for 14–20 h after transfection. Total protein extracts were collected from the protoplasts in the rice protein extraction buffer as described above. The supernatants were incubated with 10 µl Protein G Agarose beads (Millipore, 16–266) and 1 µl anti-HA (MBL, M180-3) or 1 µl anti-MYC (MBL, M192-3) antibodies at 4 °C for 5 h, and detected using anti-HA or anti-HA mAb-HRP-DirecT (MBL, M180-7, clone: TANA2; 1:1,000), anti-actin and anti-MYC or anti-MYC mAb-HRP-DirecT (MBL, M192-7, clone: My3, 1:1,000) antibodies, respectively. The intensities of protein signals were quantified using ImageJ.

For the yeast protein immunoblots, total yeast protein was extracted as previously described[51]. In brief, single yeast colonies were resuspended in 4 ml YPDA liquid medium and grown to OD_{600} = 0.6 at 30 °C. The pelleted yeast cells were resuspended in 1 ml distilled water, combined with 100 µl 0.2 M NaOH, incubated for 5 min at room temperature and pelleted and resuspended in 50 µl loading buffer. The proteins were boiled for 10 min at 95 °C and detected by immunoblotting with anti-HA or anti-MYC antibodies. Ponceau S solution was used to stain the PVDF membrane for loading control.

Rice protoplast isolation and transfection

Rice protoplasts were prepared from etiolated seedlings and transfected as previously described[11,19]. About 3–10 µg of plasmid DNA was used to transfect around 2×10^5 protoplasts using the PEG-mediated method. The fluorescent or epitope-tagged proteins were detected at 14–20 h after transfection.

Subcellular localization and BiFC analysis

For subcellular localization, the BISP–GFP plasmid and the nuclear marker bZIP63–RFP[11,19] were transformed into rice protoplasts. Following incubation at 28 °C for 14–20 h, the protoplasts were imaged using a confocal microscope (Leica, DMi8).

For BiFC analysis, protoplasts were transfected with the indicated expression vectors. Following incubation at 28 °C for 14–20 h, the protoplasts were imaged by confocal microscopy. BiFC fluorescent signals were quantified using previously described methods[52]. The intensity of YFP and CFP signals in individual protoplasts was determined using ImageJ software. The average YFP/CFP intensity ratio (per cent) of the whole cell was measured, and the means were calculated from five to ten independent cells. The fluorescent or epitope-tagged proteins were detected with anti-GFP (Roche, 11814460001; 1:1,000), anti-HA or anti-MYC antibodies.

RNA isolation, real-time and semi-quantitative RT–PCR analysis

Total RNA was isolated from rice leaf sheaths, BPHs or *N. benthamiana* leaves using TRIzol reagent (Takara, 9109). First-strand cDNA was synthesized using a PrimeScript RT Reagent kit with gDNA Eraser (Takara, RR047A) following the manufacturer's instructions.

Real-time RT–qPCR was carried out using a CFX96 Real-Time system (Bio-Rad) with SYBR Green Supermix (Bio-Rad, 172–5274) following the manufacturer's instructions. For RT–qPCR in rice, three housekeeping genes, Os*ubiquitin*, Os*tbp* and Os*hsp*[53], were used as reference genes for calibration of real-time RT–PCR data. For RT–qPCR in BPH, Nl*RPS18*, Nl*ubiquitin* and Nl*ACTB* (which encodes β-actin) were used as reference genes. The sequences of the primers used are listed in Supplementary Table 3.

For the semi-quantitative RT–PCR assays, first-strand cDNA (5 μl) was diluted to 50 μl final volume with TE buffer, and 1 μl of the dilution was used as template for PCR amplification using gene-specific primers. The

internal standard Nb*ACTB* was used. The number of cycles used for amplification with each primer pair was adjusted to ensure that the amplification was in the linear range. The PCR products were sequenced to ensure that they were derived from the targeted genes. The sequences of the primers used are listed in Supplementary Table 3.

Immunohistochemistry

For immunolocalization of BISP in rice leaf-sheath sections, *N14* plants at the two-leaf stage were infested with ten fourth- or fifth-instar nymphs for 2 days or served as non-infested controls. The outer leaf sheaths were collected, fixed, dehydrated, and embedded in paraffin (Paraplast Plus, Sigma-Aldrich, P3683) as previously described[11,17]. The sections were incubated in dimethylbenzene and an alcohol gradient to remove the paraffin and to gradually rehydrate. The sections were incubated with anti-BISP antibodies or pre-immune rabbit serum (prepared as described above, 1:100) at 4 °C in the dark overnight. The next day, the sections were washed 3 times (15 min each time) with PBST (PBS with 0.1% (v/v) Triton X-100), blocked with 5% bovine serum albumin in PBST at room temperature for 1 h, washed 3 times at 15-min intervals with PBST, and incubated with the secondary Cy3-conjugated goat anti-rabbit antibodies (Jackson ImmunoResearch, 111-165-003, 1:500) at 4 °C overnight. The sections were washed extensively with PBST at 15-min intervals and imaged under a confocal microscope (Leica, DMi8).

For immunolocalization in BPH salivary glands and guts, female or male insects were anaesthetized with carbon dioxide (CO_2). The salivary glands and guts were dissected in 0.65% NaCl, washed 3 times in PBST and fixed in 4% (v/v) formaldehyde in PBST at 4 °C overnight. The next day, the organs were extensively washed with PBST, incubated with anti-BISP antibodies, pre-immune serum or anti-NlSP1 antibodies[54] (1:100) at 4 °C in the dark overnight, and treated as described above. Nuclei were stained with 100 nM 4,6-diamidino-2-phenylindole (DAPI) at room temperature in the dark for 5 min. The salivary glands and guts were extensively washed with PBST and placed on a glass slide. Photographs were taken under a confocal microscope (Leica, DMi8). For the quantification of Cy3, the Cy3 and DAPI fluorescence signals in individual cells were quantified using ImageJ. The

average Cy3/DAPI intensity ratio of each cell was measured, and mean ratios were calculated from five to ten independent cells.

RNAi and bioassay of BPHs after dsRNA injection

A 506-bp fragment of *Bisp* was amplified with the primers listed in Supplementary Table 3, including the T7 promoter sequence in the forward primer. After cloning the *Bisp* fragment into the pMD18-T vector (Takara, 6011), dsRNA was synthesized from PCR-generated DNA templates using a MEGAscript T7 Transcription kit (Ambion, AM1354). BPH nymphs in their third, fourth or fifth stage were injected with *Bisp*-RNAi or *GFP*-RNAi using a Nanoliter 2010 injector (World Precision Instruments) as previously described[5]. The efficiency of gene silencing was determined by RT–qPCR at 1–10 days after dsRNA treatment.

After the injected BPHs were placed and reared on rice plants for 12 h, the survival rate, weight gain and honeydew excretion of *GFP*-RNAi BPHs, *Bisp*-RNAi BPHs and non-injected control insects were measured as described above. Five replicates were set up for each treatment to measure BPH survival rates, and 20 replicates were performed for each treatment group to measure weight gain and honeydew excretion.

Plant transformation

To constitutively express *Bisp* in rice plants, the BISP–MYC plasmids were transformed into *N14* or *Bph14* plants through *Agrobacterium*-mediated transformation[8]. To knockout *Os*NBR1 by genome editing, the CRISPR–Cas9 plasmids *Os*NBR1–Cas9 were introduced into *Agrobacterium tumefaciens* strain EHA105 and transformed into *Bph14* as described elsewhere[8]. Genomic DNA was extracted from these transformants, and PCR amplification was performed using primer pairs flanking the designed target site. The PCR products were sequenced to determine whether the gene editing was successful.

Quantification of free SA

Endogenous free SA levels were determined in rice leaf sheaths. Fifteen leaf sheaths from each plant or line were ground to a fine powder in liquid nitrogen as a replicate. Three biological replicates of each frozen sample were prepared. Each sample (100 mg) was added to 750 µl precooled extraction buffer (methanol:water:acetic acid, 80:19:1, v/v/v) supplemented with 3 µg naphthaleneacetic acid (Sigma-Aldrich, 35745) as an internal standard and shaken at 4 °C overnight. After centrifugation, the supernatant was filtered through a 0.22-µm nylon membrane and dried under a nitrogen flow at room temperature for at least 4 h. Finally, 200 µl methanol was added to dissolve the extracts. Endogenous free SA measurements were performed as described elsewhere[55].

Recombinant protein expression and purification in *E. coli* and BISP antibody specificity analysis

The *Os*RLCK185–MBP–His, BISP–GST–His and BISP(26–124)–GST–His constructs were expressed in *E. coli* BL21 (DE3) cells. GST-tagged or MBP-tagged recombinant proteins were purified using Glutathione Sepharose beads (GE Healthcare, 17075601) and Amylose Resin (New England Biolabs, E8022S), respectively, according to the manufacturer's instructions. Protein concentrations were determined by BCA protein assay (Beyotime, P0010S).

For the BISP antibody specificity analysis, the purified BISP–GST–His protein and total protein from whole insect and a twofold dilution series were separated on SDS–PAGE gels. BISP–GST–His was detected in immunoblots using anti-BISP and anti-His (GenScript, A00186; 1:2,000) antibodies, respectively. Ponceau S solution was used to stain the PVDF membrane and used as the total BPH protein loading control.

In vitro phosphorylation assays

To test the autophosphorylation of *Os*RLCK185, 1 µg purified recombinant *Os*RLCK185–MBP–His, BISP(26–124)–GST–His or BISP–GST–His was incubated in 25 µl kinase reaction buffer containing 50 mM HEPES (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, and 2.5 µCi [γ-^{32}P]ATP at 30 °C for 10–30 min. The kinase reaction was subsequently analysed using a

10% SDS–PAGE gel for autoradiography. A similar gel was subjected to Coomassie Brilliant Blue staining as a control.

To test the phosphorylation of *Os*RLCK185 at the serine and threonine residues, phosphorylated *Os*RLCK185 was detected by immunoblotting with anti-phosphoserine/phosphothreonine antibodies (Millipore, 05-368, 1:1000) and anti-His antibody (GenScript, A00186).

Recombinant protein expression and purification in insect cells and BLI analyses

LRR, N14–LRR, BISP, BISP(26–124), and BISP(125–241) were cloned into pFastBac 1 with an N-terminal 6×His-SUMO tag and expressed in SF9 insect cells (Thermo Fisher Scientific, 11496015) at 27 °C. SF9 cells (500 ml) cultured in Sf-900 II SFM medium (Invitrogen, 10902088) were infected with 25 ml recombinant baculovirus and were cultured for 4 days at 27 °C. Cells were collected and re-suspended in PBS solution (2 mM KH_2PO_4, 8 mM Na_2HPO_4, 136 mM NaCl, and 2.6 mM KCl, pH 7.4). After sonication and centrifugation, recombinant LRR, N14–LRR, BISP, BISP(26–124), and BISP(125–241) proteins in the supernatant were purified with Ni-NTA beads (Novagen,70666-4) and then were dialysed in PBS solution.

The binding kinetics and affinities of LRR and N14–LRR with BISP, BISP(26–124) or BISP(125–241) was determined using an Octet RED96 system (FortéBio) as previously described[56,57]. Purified LRR and N14–LRR were biotinylated with NHS-biotin (Sangon Biotech, C100212) according to the manufacturer's instructions. All streptavidin-coated biosensors (FortéBio) were hydrated in PBS buffer (2 mM KH_2PO_4, 8 mM Na_2HPO_4, 136 mM NaCl, 2.6 mM KCl, 0.05% Tween 20, and 0.5% bovine serum albumin, pH 7.4) for 10 min. Biotinylated LRR and N14–LRR were diluted in PBS buffer to a final concentration of 20 $\mu g\ ml^{-1}$ and immobilized onto a streptavidin-coated biosensor. Biosensors with immobilized LRR and N14–LRR were diluted in binding buffer containing different concentrations of BISP, BISP(26–124) or BISP(125–241) for the binding kinetics analysis (association step). Subsequently, the streptavidin-coated biosensor was returned to PBS buffer for 180 s (dissociate step). PBS buffer was used in all

BLI experiments to reduce nonspecific binding to the biosensors except for experiments in the presence of BISP, BISP(26–124) or BISP(125–241). The experiments included the following steps at 25 °C: (1) loading (60 s); (2) baseline (120 s); (3) immobilization of proteins onto sensors (60 s); (4) association (180 s); and (5) disassociation (180 s). The data were analysed, and sensor-grams were step-corrected and reference-corrected. Global fitting of the kinetic rates was performed, then the equilibrium dissociation constant (K_d), association (K_{on}) and dissociation (K_{off}) rate constants and their standard errors were analysed using FortéBio data analysis software (v.1.1.0.16, FortéBio). The assays were repeated two times for each affinity measurement.

For the competition assay, an Octet RED96 system (FortéBio) was used as previously described[58]. Biotinylated LRR was diluted in PBS buffer to a final concentration of 20 μg ml^{-1} and immobilized onto a streptavidin-coated biosensor. The association of BISP(26–124) or PBS was measured for 360 s at 25 °C. Next, the competitors (BISP, PBS or BISP(26–124)) were added to the BISP(26–124)-associated or PBS-associated samples. The experiments included the following steps at 25 °C: (1) loading (60 s); (2) baseline (120 s); (3) immobilization of protein onto sensors (60 s); (4) baseline (240 s); (5) association with BISP(26–124) or PBS (360 s); (6) disassociation (240 s); (7) association with the competitors BISP, PBS or BISP(26–124) (120 s); and (8) dissociation (180 s). The data were analysed using FortéBio data analysis software (v.1.1.0.16, FortéBio). The assays were repeated two times for each measurement.

MST assay

The MST assay was performed as previously described[59,60]. The affinity of purified LRR or N14–LRR with BISP, BISP(26–124) or BISP(125–241) was measured using a Monolith NT.115 (Nanotemper Technologies). Proteins were fluorescently labelled using a Monolith Protein Labelling kit (RED-NHS 2nd Generation; Nanotemper Technologies, MO-L011) according to the manufacturer's protocol, and the labelled protein used for each assay was about 20 nM. A solution of unlabelled BISP, BISP(26–124) or BISP(125–241) protein was diluted for an appropriate serial concentration gradient. The samples were loaded into capillaries (Nanotemper Technologies, MO-

K022) after incubation at room temperature for 5 min. Measurements were performed in 20 µl of PBS buffer containing 0.05% Tween 20. The assays were repeated three times for each affinity measurement. Data analyses were performed using Nanotemper Analysis software provided by the manufacturer.

For the competition assays between LRR with BISP and BISP(26–124) by MST, 50 nM LRR proteins and 200 nM BISP(26–124) were combined in PBS buffer with 0.05% Tween 20 and incubated for 10 min to enable protein interactions. A series of different concentrations of BISP was then added to the LRR–BISP(26–124) mixture before measurement. The assays were repeated three times for each measurement. The dissociation constant (K_i) from the MST assays was calculated using the K_i Finder software from the Nanotemper website (https://nanotemper.my.site.com/explore/s/article/Can-competition-assays-be-performed-with-MST).

Inhibition of protein degradation in rice protoplasts

After PEG-mediated transformation, the rice protoplasts were placed in 1 ml W5 buffer containing 10 µM 3-MA (3-methyladenine, Selleck, S2767), 100 µM chloroquine (MCE, HY-17589A), 100 µM leupeptin (MCE, HY-18234A), or 20 µM E-64d (Aloxistatin, Sigma-Aldrich, E8640) to inhibit autophagy or 50 µM MG132 (Selleck, S2619) to inhibit the 26S proteasome. The treated protoplasts were incubated at 28 °C for 14–20 h before protein extraction. Anti-actin (Abbkine, ABL1050) was used as a reference to quantify total protein levels. Rice WRKY72 is degraded by the 26S proteasome pathway[19] and was used as a positive control to ensure that the proteasome inhibitor MG132 was active. The rice HD1 protein is degraded in the dark by the autophagy pathway[61], and *Os*NBR1 was used as a positive control to assure autophagy inhibitors were active. The intensity of protein signals in immunoblots was quantified by ImageJ.

Live-cell imaging of autophagic structures and virus-induced gene silencing in *N. benthamiana* leaves

To observe autophagic structures in *N. benthamiana* leaves by confocal microscopy, *A. tumefaciens* strain GV3101 cells harbouring CFP–*Nb*ATG8F

(designed as CFP–ATG8F)[27] and the constructed expression vectors, BISP-201, BPH14-203, N14-203, *Os*NBR1-203, and N1-203 were infiltrated into *N. benthamiana* leaves for 60 h. The samples were then infiltrated with or without the autophagy inhibitor ConA (first made in DMSO at 0.5 M, then 1:1,000 added) for 10 h before being examined under a confocal microscope (Leica, DMi8).

For virus-induced gene silencing (VIGS) assays, pTRV1 and pTRV2 or its derivatives were introduced into *Agrobacterium* strain GV3101. VIGS assays were performed as previously described[62]. In brief, the *Agrobacterium* cultures with pTRV1 and pTRV2 or its derivatives were mixed at a 1:1 ratio and were infiltrated into the leaves of six-leaf stage *N. benthamiana* plants. Silenced phenotypes appeared in the upper leaves about 10 days after infiltration. *A. tumefaciens* strain GV3101 cells harbouring CFP–*Nb*ATG8F and the constructed expression vectors BISP-101 (BISP–YFP) and BPH14-203 were infiltrated into the *N. benthamiana* upper leaves.

Transmission electron microscopy

For immunogold electron microscopy, *N14* leaf sheaths infested with BPH for 24 h were cut into 0.1-mm pieces, fixed in precooled 3% paraformaldehyde with 0.1% glutaraldehyde (in 100 mM PBS, pH 7.4) and immediately placed under a vacuum until the samples were completely submerged in the liquid. The samples were transferred to fresh 0.1% glutaraldehyde and incubated at 4 °C overnight. The next day, the samples were washed 3 times with PBS (pH 7.4) for 30 min. The samples were cleared by successive washes with 30% and 50% ethanol for 60 min each at 4 °C, then 70%, 80%, 85%, 90%, and 95% ethanol for 30 min each at −20 °C, and finally 100% ethanol 3 times for 60 min each at −20 °C. After ethanol washes, the samples were embedded in Lowicry K4M resin stepwise, according to the manufacturer's instructions. Ultrathin sections (70 nm) of the embedded tissues were cut with a diamond knife using an Ultracut E Ultramicrotome (Reichart-Jung) and collected on 3-mm copper (mesh) grids. Immunolabelling of leaf sections was done using standard procedures as previously described[59] with anti-BISP antibodies or pre-immune serum at 100 μg ml^{-1} and anti-rabbit IgG gold-coupled secondary antibodies (Sigma-Aldrich, G7277- 4ML) at 1:50 dilution. The sections

were then stained with 2% uranyl acetate (in 70% ethanol) for 15 min, washed in distilled water 3 times, stained with lead citrate for 12 min, washed in NaOH buffer, washed in distilled water 3 times, and examined under a JEM-1230 electron microscope.

For transmission electron microscopy observations, 2-week-old *Bph14–Bisp* transgenic lines, *Bph14* and MH63 leaf sheaths were cut into 0.1-mm pieces, fixed in precooled 4% paraformaldehyde with 2% glutaraldehyde (in 100 mM PBS, pH 7.4) and immediately placed under a vacuum until the samples were completely submerged in the liquid. The samples were transferred to fresh 2% glutaraldehyde and incubated at 4 °C overnight. The next day, the samples were washed 5 times with PBS (pH 7.4) for 20 min and submerged in a 1% solution of osmium tetroxide at room temperature for 1–2 h until they turned completely black because of oxidation. The samples were washed 5 times with PBS at 20-min intervals. The samples were cleared by successive washes with 15%, 30%, 50%, and 70% ethanol for 30 min each, then 80%, 85%, 90%, and 95% ethanol for 20 min each, and finally 100% ethanol twice for 45 min each. After ethanol washing, the samples were embedded in Spurr resin stepwise as instructed by the manufacturer. Ultrathin sections (70 nm) of the embedded tissues were cut with a diamond knife using an Ultracut E Ultramicrotome (Reichart-Jung), collected on 3-mm copper (mesh) grids, stained with 2% uranyl acetate (in 70% ethanol) for 10 min, washed in distilled water 3 times, stained with lead citrate for 15 min, washed in NaOH buffer, and examined under a JEM-1230 electron microscope. To quantify autophagosomes, the double-membrane autophagosomes in every eight transmission electron microscope images were counted and served as one biological replicate; each genotype had five replicates.

BPH infestation

Seeds were sown in 10-cm diameter plastic cups and grown in a plant growth chamber (CONVIRON PGC2000) as described elsewhere[11]. To analyse the relationship between BISP levels and autophagy, 4-leaf rice plants were infested with second- or third-instar BPH nymphs (10 per plant) at the selected time points (6, 12, 24, 48, and 72 h before the end of the experiment). Control groups were maintained in parallel but without BPH

infestation. At the end of the treatments, leaf sheaths were wiped with cotton wool balls soaked in alcohol to remove honeydew, leaf sheaths were excised, frozen, and ground to a fine powder in liquid nitrogen, and proteins isolated. Immunoblotting was performed following the procedures described above in the section 'Protein extraction, immunoblotting and in vivo co-IP assays'. All treatments were repeated three times.

To measure autophagy and BISP levels following BPH infestation, 4-leaf rice plants were infested with second- or third-instar BPH nymphs (10 per plant) for 24 h. The insects were then removed, and the plants were allowed to grow under standard conditions (described above). At the selected time points after BPH removal, honeydew was removed from leaf sheaths, leaves were excised, frozen, and ground to a fine powder in liquid nitrogen, and proteins isolated. Immunoblotting was performed following the procedures described above. All treatments were repeated three times.

For the autophagy inhibitor ConA treatment, the BPH-infested and non-infested rice plants were treated with or without ConA (10 μM) for 12 h before samples were collected. Leaf sheath total protein extraction and immunoblotting were performed according to the procedures described above. All treatments were repeated three times.

Agronomic performance test

To investigate the agronomic performance of *Bph14* and *Bph14–Bisp* plants, plants were grown in a field in Wuhan, China, under routine management. At harvest, the middle plants in the central row of each plot were assessed for the following nine agronomic traits: plant height, sword leaf length, sword leaf width, number of panicles per plant, number of filled grains per panicle, number of filled grains per plant, percentage of filled grain, 1,000-grain weight, and grain yield per plant[63].

Statistical analysis

No statistical methods were used to predetermine the sample sizes. The experiments were not randomized, and investigators were not blinded to allocation during the experiments and outcome assessment. Quantification

analyses of protein abundance were conducted using ImageJ software. All values are presented as the mean ± s.d. or mean ± s.e.m. as indicated. Data points are plotted on the graphs, and the number of samples is indicated in the corresponding figure legends. In all cases, sampling for replicates and between time points was performed by collecting material from new plants. All statistical analyses were performed by one-way analysis of variance with MS Excel and GraphPad Prism (v.8.0.1, GraphPad Software). Detailed information about statistical analysis values for all experiments is provided in Supplementary Table 4.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data are available within this article and its Supplementary Information. Original gel blots are shown in Supplementary Fig. 1. Original data points in graphs are shown in the source data files. Statistical analyses of this study are provided in Supplementary Table 4. The sequences of *Bisp* have been deposited and made publicly available in GenBank with accession number MH885414. Source data are provided with this paper.

References

1. Kaloshian, I. & Walling, L. Hemipteran and dipteran pests: effectors and plant host immune regulators. *J. Integr. Plant Biol.* **58**, 350–361 (2016).

2. Hatchett, J. H. & Gallun, R. L. Genetics of the ability of the Hessian fly, *Mayetiola destructor*, to survive on wheats having different genes for resistance. *Ann. Entomol. Soc. Am.* **63**, 1400–1407 (1970).

3. Oerke, E. C. Crop losses to pests. *J. Agric. Sci.* **144**, 31–43 (2006).

4. Jiang, Y., Zhang, C., Chen, R. & He, S. Challenging battles of plants with phloem-feeding insects and prokaryotic pathogens. *Proc. Natl Acad. Sci. USA* **116**, 23390–23397 (2019).

5. Shangguan, X. et al. A mucin-like protein of planthopper is required for feeding and induces immunity response in plants. *Plant Physiol.* **176**, 552–565 (2018).

6. Rossi, M. et al. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proc. Natl Acad. Sci. USA* **95**, 9750–9754 (1998).

7. Dogimont, C., Chovelon, V., Pauquet, J., Boualem, A. & Bendahmane, A. The *Vat* locus encodes for a CC-NBS-LRR protein that confers resistance to *Aphis gossypii* infestation and *A. gossypii*-mediated virus resistance. *Plant J.* **80**, 993–1004 (2014).

8. Du, B. et al. Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proc. Natl Acad. Sci. USA* **106**, 22163–22168 (2009).

9. Tamura, Y. et al. Map-based cloning and characterization of a brown planthopper resistance gene *BPH26* from *Oryza sativa* L. ssp. *indica* cultivar ADR52. *Sci. Rep.* **4**, 5872 (2014).

10. Zhao, Y. et al. Allelic diversity in an NLR gene *BPH9* enables rice to combat planthopper variation. *Proc. Natl Acad. Sci. USA* **113**, 12850–12855 (2016).

11. Guo, J. et al. *Bph6* encodes an exocyst-localized protein and confers broad resistance to planthoppers in rice. *Nat. Genet.* **50**, 297–306 (2018).

12. Shi, S. et al. *Bph30* confers resistance to brown planthopper by fortifying sclerenchyma in rice leaf sheaths. *Mol. Plant* **14**, 1714–1732 (2021).

13. Jones, D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).

14. Cui, H., Tsuda, K. & Parker, J. E. Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* **66**, 487–511 (2015).

15. Dyck, V. & Thomas, B. in *Brown Planthopper: Threat to Rice Production in Asia* (ed. Brady, N. C.) 3–17 (International Rice Research Institute, 1979).

16. Sogawa, K. The rice brown planthopper: feeding physiology and host plant interactions. *Annu. Rev. Entomol.* **27**, 49–73 (1982).

17. Hao, P. et al. Herbivore-induced callose deposition on the sieve plates of rice: an important mechanism for host resistance. *Plant Physiol.* **146**, 1810–1820 (2008).

18. Liu, F. et al. *Ultrabithorax* is a key regulator for the dimorphism of wings, a main cause for the outbreak of planthoppers in rice. *Nat. Sci. Rev.* **7**, 1181–1189 (2020).

19. Hu, L. et al. The coiled-coil and nucleotide binding domains of Brown Planthopper Resistance14 function in signaling and resistance against planthopper in rice. *Plant Cell* **29**, 3157–3185 (2017).

20. Tang, D., Wang, G. & Zhou, J. Receptor kinases in plant–pathogen interactions: more than pattern recognition. *Plant Cell* **29**, 618–637 (2017).

21. Yamaguchi, K. et al. A receptor-like cytoplasmic kinase targeted by a plant pathogen effector is directly phosphorylated by the chitin receptor and mediates rice immunity. *Cell Host Microbe* **13**, 347–357 (2013).

22. Wang, C. et al. OsCERK1-mediated chitin perception and immune signaling requires receptor-like cytoplasmic kinase 185 to activate an MAPK cascade in rice. *Mol. Plant* **10**, 619–633 (2017).

23. Nelson, R., Wiesner-Hanks, T., Wisser, R. & Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **19**, 21–33 (2018).

24. Xu, G. et al. uORF-mediated translation allows engineered plant disease resistance without fitness costs. *Nature* **545**, 491–494 (2017).

25. Dikic, I. Proteasomal and autophagic degradation systems. *Annu. Rev. Biochem.* **86**, 193–224 (2017).

26. Wersch, S., Tian, L., Hoy, R. & Li, X. Plant NLRs: the whistleblowers of plant immunity. *Plant Commun.* **1**, 100016 (2020).

27. Han, S. et al. Cytoplastic glyceraldehyde-3-phosphate dehydrogenases interact with ATG3 to negatively regulate autophagy and immunity in *Nicotiana benthamiana*. *Plant Cell* **27**, 1316–1331 (2015).

28. Klionsky, D. et al. Guidelines for the use and interpretation of assays for monitoring autophagy (4th edition). *Autophagy* **17**, 1–382 (2021).

29. Dikic, I. & Elazar, Z. Mechanism and medical implications of mammalian autophagy. *Nat. Rev. Mol. Cell Biol.* **19**, 349–364 (2018).

30. Kellner, R., Concepcion, J., Maqbool, A., Kamoun, S. & Dagdas, Y. ATG8 expansion: a driver of selective autophagy diversification? *Trends Plant Sci.* **22**, 204–214 (2017).

31. Yu, J., Zhen, X., Li, X., Li, N. & Xu, F. Increased autophagy of rice can increase yield and nitrogen use efficiency (NUE). *Front. Plant Sci.* **10**, 584 (2019).

32. Brillada, C. et al. Exocyst subunit Exo70B2 is linked to immune signaling and autophagy. *Plant Cell* **33**, 404–419 (2021).

33. Zhang, Y. & Chen, Z. Broad and complex roles of NBR1-mediated selective autophagy in plant stress responses. *Cells* **9**, 2562 (2020).

34. Hofius, D., Li, L., Hafrén, A. & Coll, N. Autophagy as an emerging arena for plant–pathogen interactions. *Curr. Opin. Plant Biol.* **38**, 117–123 (2017).

35. Thirumalaikumar, V. et al. Selective autophagy regulates heat stress memory in *Arabidopsis* by NBR1-mediated targeting of HSP90 and

ROF1. *Autophagy* **17**, 2184–2199 (2021).

36. Hafrén, A. et al. Selective autophagy limits cauliflower mosaic virus infection by NBR1-mediated targeting of viral capsid protein and particles. *Proc. Natl Acad. Sci. USA* **114**, E2026–E2035 (2017).

37. Svenning, S., Lamark, T., Krause, K. & Johansen, T. Plant NBR1 is a selective autophagy substrate and a functional hybrid of the mammalian autophagic adapters NBR1 and p62/SQSTM1. *Autophagy* **7**, 993–1010 (2011).

38. Will, T. & Bel, A. Physical and chemical interactions between aphids and plants. *J. Exp. Bot.* **57**, 729–737 (2006).

39. Ghaffar, M., Pritchard, J. & Ford-Lioyd, B. Brown planthopper (*N. lugens* Stål) feeding behaviour on rice germplasm as an indicator of resistance. *PLoS ONE* **6**, e22137 (2011).

40. Michel, A. & Harris, M. Editorial overview: why modern research justifies the re-emergence of host–plant resistance as a focus for pest management. *Curr. Opin. Insect Sci.* **45**, iii–v (2021).

41. Wang, J. et al. Two VOZ transcription factors link an E3 ligase and an NLR immune receptor to modulate immunity in rice. *Mol. Plant* **14**, 253–266 (2021).

42. Rodriguez, E. et al. Autophagy mediates temporary reprogramming and dedifferentiation in plant somatic cells. *EMBO J.* **39**, e103315 (2020).

43. Jiang, H. et al. Evaluation and breeding application of six brown planthopper resistance genes in rice maintainer line Jin 23B. *Rice* **11**, 22 (2018).

44. Shi, Z. Y. et al. Construction of genomic library of a BPH-resistant rice line with binary vector and physical map of *Qbp1* locus. *Plant Sci.* **165**, 879–885 (2003).

45. Lu, Y. et al. Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. *Mol. Plant* **10**, 1242–1245 (2017).

46. Huang, Z. et al. Identification and mapping of two brown planthopper resistance genes in rice. *Theor. Appl. Genet.* **102**, 929–934 (2001).

47. Heinrichs, E., Medrano, F. & Rapusas, H. *Genetic Evaluation for Insect Resistance in Rice* (International Rice Research Institute, 1985).

48. Earley, K. et al. Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616–629 (2006).

49. Ma, X. et al. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* **8**, 1274–1284 (2015).

50. Hofius, D. et al. Autophagic components contribute to hypersensitive cell death in *Arabidopsis*. *Cell* **137**, 773–783 (2009).

51. Kushnirov, V. Rapid and reliable protein extraction from yeast. *Yeast* **16**, 857–860 (2000).

52. Kudal, J. & Bock, R. Lighting the way to protein–protein interactions: recommendations on best practices for bimolecular fluorescence complementation analyses. *Plant Cell* **28**, 1002–1008 (2016).

53. Hu, J. et al. The *Bphi008a* gene interacts with the ethylene pathway and transcriptionally regulates MAPK genes in the response of rice to brown planthopper feeding. *Plant Physiol.* **156**, 856–872 (2011).

54. Huang, J. et al. Salivary Protein 1 of brown planthopper is required for survival and induces immunity response in plants. *Front. Plant Sci.* **11**, 571280 (2020).

55. Liu, H., Li, X., Xiao, J. & Wang, S. A convenient method for simultaneous quantification of multiple phytohormones and metabolites: application in study of rice–bacterium interaction. *Plant Methods* **8**, 2 (2012).

56. Cao, L. et al. Design of protein-binding proteins from the target structure alone. *Nature* **605**, 551–560 (2022).

57. Hong, Q. et al. Molecular basis of receptor binding and antibody neutralization of omicron. *Nature* **604**, 546–552 (2022).

58. Alanine, D. et al. Human antibodies that slow erythrocyte invasion potentiate malaria-neutralizing antibodies. *Cell* **178**, 216–288 (2019).

59. Friml, J. et al. ABP1-TMK auxin perception for global phosphorylation and auxin canalization. *Nature* **609**, 575–581 (2022).

60. Mamou, G. et al. Peptidoglycan maturation controls outer membrane protein assembly. *Nature* **606**, 953–959 (2022).

61. Hu, Z. et al. Autophagy targets Hd1 for vacuolar degradation to regulate rice flowering. *Mol. Plant* **15**, 1137–1156 (2022).

62. Liu, Y. et al. Autophagy regulates programmed cell death during the plant innate immune response. *Cell* **121**, 567–577 (2005).

63. You, A. Q. et al. Identification of QTLs across recombinant inbred lines and testcross populations for traits of agronomic importance in rice. *Genetics* **172**, 1287–1300 (2006).

Acknowledgements

We thank Y. Liu (Tsinghua University) for providing plasmids CFP–*Nb*ATG8F, pTRV1, pTRV2, pTRV2-*Nb*PI3K, pTRV2-*Nb*ATG6 and pTRV2-*Nb*ATG7; Y. Liu (South China Agricultural University) for providing plasmid pYLCRISPR/Cas9-MH; S. Zhang (Huazhong Agricultural University) for helpful suggestions for the BLI assays; and J. L. Dangl, the University of North Carolina, for critical reading and insightful discussions. This work was supported by grants from the National Natural Science Foundation of China (31630063, 31901518 and 31230060), the Major Program of Guangdong Basic and Applied Research, China (2019B030302006), Program of Department of Science and Technology of Hubei Province (2022EJD011), the National Key Research and Development Program of China (2016YFD0100903 and 2016YFD0100600), the National Program on Research & Development of Transgenic Plants (2016ZX08009-003-001), the National Postdoctoral

Program for Innovative Talents (BX20180230) and the China Postdoctoral Science Foundation (2018M640729).

Author information

Author notes

1. These authors contributed equally: Jianping Guo, Huiying Wang, Wei Guan

Authors and Affiliations

1. State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China

   Jianping Guo, Huiying Wang, Wei Guan, Qin Guo, Jing Wang, Jing Yang, Yaxin Peng, Junhan Shan, Mingyang Gao, Shaojie Shi, Xinxin Shangguan, Bingfang Liu, Shengli Jing, Jing Zhang, Chunxue Xu, Jin Huang, Weiwei Rao, Xiaohong Zheng, Di Wu, Cong Zhou, Bo Du, Rongzhi Chen, Lili Zhu, Yuxian Zhu & Guangcun He

2. Hubei Hongshan Laboratory, Wuhan, China

   Jianping Guo, Wei Guan, Yuxian Zhu, Qifa Zhang & Guangcun He

3. Key Laboratory of Biology and Genetic Improvement of Oil Crops, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China

   Huiying Wang

4. Institute for Advanced Studies, Wuhan University, Wuhan, China

   Yuxian Zhu

5. Department of Botany and Plant Sciences, University of California, Riverside, CA, USA

Linda L. Walling

6. National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China

   Qifa Zhang

Contributions

G.H. conceived and supervised the project. J.G., H.W., W.G. and G.H. designed the experiments. J.G., H.W. and W.G. carried out most of the experiments, including the Y2H assays, cloning, co-IP, BiFC, subcellular localization, autophagy analysis, BLI, MST, phenotype assays, microscopy, RNA analysis and western blotting. Q.G., J.W. and J.S. performed transmission electron microscopy, SA measurement and phosphorylation assays. J.Y., S.S., X.S., S.J. and B.L. performed cloning, dsRNA injection, BPH performance analysis and RNA analysis. Y.P., M.G., X.Z. and D.W. performed cloning, western blotting, protoplast isolation and immunogold microscopy. J.Z. and C.X. performed VIGS and microscopy. J.H., W.R. and C.Z. maintained the BPH population, performed phenotype and immunohistochemistry. B.D., R.C. and L.Z generated material used in this study and performed agronomic traits. J.G., H.W. and W.G. prepared the figures. Y.Z., Q.Z. and L.L.W. provided theoretical contributions to the project. J.G., H.W., W.G., Y.Z., Q.Z., L.L.W. and G.H. analysed data and wrote the manuscript.

Corresponding author

Correspondence to Guangcun He.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Thorsten Nurnberger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Host-plant resistance evaluation, BISP subcellular localization in rice protoplasts, *Bisp* gene and BISP protein expression profile in BPH, and BISP delivery into rice leaf sheath.

a, BPH-resistance scores of *Bph14* and *N14* plants after BPH infestation. *Bph14* and *N14* plants express the resistant NLR (BPH14) and the susceptible allele variant (N14), respectively. Lower scores indicate higher resistance to the insects. Data were collected 7 days after BPH infestation ($n = 36$ plants examined over 3 independent experiments). *P* values were derived by one-way ANOVA. **b**, Honeydew excretion of BPHs fed on *Bph14* and *N14* plants for 2 days. *P* values were derived by one-way ANOVA ($n = 20$, biologically independent plants). **c**, BISP-GFP had a

nucleocytoplasmic localization in rice protoplasts. bZIP63-RFP was used as a nuclear marker. The confocal images were taken 14–20 h after transformation. Scale bar, 5 μm. **d**, Immunoblot detection of full-length BISP-GFP expression in rice protoplasts. BISP-GFP was detected by anti-GFP antibody. **e**, Immunoblot detection of the expressed proteins in BiFC assays. BISP-YN, $BISP^{125–241}$-YN(BISP(125–241)-YN), and BPH14-YC or N14-YC were co-expressed in rice protoplasts and the total proteins were detected using anti-Myc, anti-HA or anti-ACTIN antibody. **f**, Relative levels of *Bisp* RNA in different BPH tissues. SG, salivary glands; FB, fat body. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **g**, Relative level of *Bisp* RNAs during different developmental stages. 1st to 5th, are first to fifth instars; Female and Male are adults. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **h**, (*Upper*) Determination of the anti-BISP specificity using purified BISP-GST-His. The purified BISP-GST-His protein was serially diluted and detected by anti-BISP or anti-His antisera. (*Bottom*) Correlation between the anti-BISP labeled signal and anti-His labeled signals. The densities of anti-BISP and anti-His labeled signals were quantified using ImageJ with the undiluted sample set as 1. **i**, (*Upper*) Determination of the anti-BISP specificity using a dilution series of total mixed adults BPH proteins. BISP was detected using the anti-BISP antiserum. Ponceau staining (Ponceau S) is a loading control. (*Bottom*) Correlation between the anti-BISP labeled and Ponceau S signals. The density of anti-BISP labeled and Ponceau S signals were quantified using ImageJ with the undiluted sample set as 1. **j**,**k**, Immunohistochemical localization of BISP in male BPH salivary glands. Anti-BISP (**j**) or pre-immune rabbit serum (**k**) was used as the primary antiserum. Red fluorescence (Cy3) shows the localization of BISP and blue fluorescence shows the location of DAPI-stained nuclei. Scale bars, 100 μm. **l**,**m**, Immunohistochemical localization of BISP in mixed adults BPH guts. Anti-BISP (**l**) or pre-immune rabbit serum (**m**) was used as the primary antiserum. Red fluorescence (Cy3) shows the localization of BISP and blue fluorescence shows the location of DAPI-stained nuclei. Scale bars, 100 μm. **n**,**o**, Immunohistochemical localization of *Nl*SP1 in female BPH salivary glands. Either anti-*Nl*SP1 (**n**) or pre-immune rabbit serum (**o**) was used as the primary antiserum. Red fluorescence (Cy3) shows the localization of NlSP1 and blue fluorescence shows the location of DAPI-stained nuclei. Scale bars, 100 μm. **p**, Quantification of Cy3 intensities. The

Cy3 fluorescence in samples using anti-BISP, anti-*Nl*SP1 and pre-immune serum was normalized relative to DAPI fluorescence (blue) intensity. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 5$, biologically independent experiments). **q-t**, Immunogold electron microscopy analysis of the distribution of BISP in the rice leaf sheaths during BPH feeding. **q**,**s**, TEM image of the BPH salivary sheaths in the rice phloem tissue. Either anti-BISP (**q**) or pre-immune rabbit serum (**s**) was used as the primary antiserum. Scale bars, 20 µm. **r**,**t**, TEM image of the boxed area in (**q**) and (**s**). Red arrows indicate the anti-BISP-labeled immunogold particles (dark spots). Scale bars, 200 nm. SH, salivary sheaths; P, phloem cells. In box plots in **a**,**b**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted. Experiments were independently repeated two (**q**–**t**) or three (**b**–**e**, **h**–**o**) times, each giving similar results.

Source data

Extended Data Fig. 2 BISP played crucial roles in the feeding and performance of BPH on rice plants.

a, Effects of dsRNA injection on *Bisp* transcript levels in BPHs. *Bisp*-Ri (*Bisp*-RNAi) insects were assayed daily for 10 days. *GFP*-Ri (*GFP*-RNAi) and CK insects were assayed on day 3. Data are mean ± s.d. ($n = 3$, biologically independent experiments). The inset shows a representative immunoblot demonstrating the relative abundance of BISP in *GFP*-Ri BPHs and *Bisp*-Ri BPHs at 3 days after injection. Total BPH proteins were extracted from injected or control insects and immunoblotted. BISP was detected with anti-BISP antibodies and Ponceau S staining was used as the loading control. **b**,**c**,**d**, Weight gain (**b**), honeydew excretion (**c**) and survival rates (**d**) of dsRNA-injected and non-injected BPHs fed on *N14* plants. **e**,**f**,**g**, Weight gain (**e**), honeydew excretion (**f**) and survival rates (**g**) of dsRNA-injected and non-injected BPHs fed on the *Osrlck185*$^{-}$-1 (Os*rlck185*-1) and WT plants. **h**,**i**,**j**, Weight gain (**h**), honeydew excretion (**i**) and survival rates (**j**) of dsRNA-injected BPHs fed on the *Bph14* and MH63 plants. In **a**–**j**, BPHs were microinjected with *Bisp* dsRNA (*Bisp*-Ri) or *GFP* dsRNA (*GFP*-Ri) and non-injected BPHs (CK)

served as a control. For **b**–**j**, *P* values were derived by one-way ANOVA. In **b**,**c**,**e**,**f**,**h**,**i**, data are mean ± s.d. ($n = 20$, biologically independent plants); in **d**,**g**,**j**, data are mean ± s.e.m. ($n = 5$, biologically independent experiments). In box plots in **b**,**c**,**e**,**f**,**h**,**i**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted. In **e**–**j**, different letters indicate significant differences ($P < 0.05$, one-way ANOVA); exact P values are provided in Supplementary Table 4. All the experiments were repeated three times, each giving similar results.

Source data

Extended Data Fig. 3 Effects of *Bisp* on BPH resistance in *N14-Bisp* transgenic lines.

a, Photographs of the independent *N14-Bisp* homozygous T_2 transgenic lines (*N14-Bisp*-3 and *N14-Bisp*-11) and *N14* plants at the 3-leaf stage without BPH feeding. Scale bar, 10 cm. **b**, Plant height of *N14-Bisp* lines and *N14* at the 3-leaf stage. *P* values were derived by one-way ANOVA ($n = 30$, biologically independent samples). **c**, **d**, Weight gain (**c**) and honeydew excretion (**d**) of BPHs reared on the *N14-Bisp* transgenic and *N14* plants for 2 d. *P* values were derived by one-way ANOVA ($n = 10$, biologically independent samples). **e**, Results of the two-host choice assay examining the settling of BPHs on the *N14-Bisp* transgenic and *N14* plants. *P* values were derived by one-way ANOVA. Data are means ± s.d. (for *N14-3* and *N14-Bisp*-11, $n = 7$ and 10, respectively, biologically independent experiments). **f**, Transcript levels analysis of SA biosynthesis- and signaling-related genes *OsICS1* and *OsNPR1* and defense-related genes *OsPR1a*, *OsPR1b*, *OsPR5* and *OsPR10* in two *N14-Bisp* transgenic lines and *N14* plants. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). In box plots in **b**, **c**, **d**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted. All the experiments were repeated three times, each giving similar results.

Source data

[Extended Data Fig. 4 Identification of OsRLCK185 and its role in plant defense.]

a, Y2H screen for BISP-interacting kinases. Yeast strain AH109 was co-transformed with the indicated constructs and grown on DDO (SD/-Leu-Trp) and TDO (SD/-Leu-Trp-His) selective medium containing 1.5 mM 3-amino-1,2,4-triazole (3-AT). **b**, Immunoblot detection of proteins expressed in yeast using anti-HA and anti-Myc (anti-MYC) antibodies, respectively. Ponceau S staining was used as the loading control. Asterisks indicate specific bands detected by immunoblotting analysis. **c**, Structure of *Os*RLCK185. The orange box indicates the kinase domain. **d**, BISP but not $BISP^{26–124}$ (BISP(26–124)) reduced OsRLCK185 autophosphorylation activity at Ser and Thr residues *in vitro*. OsRLCK185 proteins were immunoblotted and phosphorylated OsRLCK185 was detected with anti-phosphoserine/ phosphothreonine antibodies. The input proteins were immunoblotted and detected with anti-His antibodies. **e**, Verification of three independent Os*RLCK185* knockout lines (Os*rlck185*, *Osrlck185*$^{−}$-1, −2, and −3) by PCR-based sequencing. **f**,**g**, Weight gain (**f**) and honeydew excretion (**g**) of BPHs on three *Osrlck185*$^{−}$ lines and WT plants for 2 d. *P* values were derived by one-way ANOVA ($n = 30$, biologically independent samples). In box plots, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the centre line indicates the median. Individual data points are plotted. Experiments were repeated three times (**a**,**b**,**d**,**f**,**g**) or two times (**e**), each giving similar results.

[Source data]

[Extended Data Fig. 5 BISP triggered *Bph14*-mediated BPH resistance in *Bph14* plants.]

a, Y2H assays between the full-length BPH14, the BPH14 LRR domain (LRR) and the N14 LRR domain (N14-LRR) and the BISP deletion mutants $BISP^{26–124}$ (BISP(26–124)) and $BISP^{125–241}$(BISP(125–241)). DDO, SD/-Leu-Trp; TDO, SD/-Leu-Trp-His with 2.5 mM 3-AT. **b**, Biolayer interferometry (BLI) determination of the binding kinetics between N14-LRR and BISP. BLI response profile of BISP at different concentrations

with the sensor-immobilized N14-LRR. **c**, BLI analysis for the binding kinetics between the BPH14 LRR and $BISP^{26–124}$. BLI response profile of $BISP^{26–124}$ at different concentrations with the sensor-immobilized LRR. K_d, equilibrium dissociation constant; K_{on}, association rate constant; K_{off}, dissociation rate constant. **d**, BLI analysis for the binding kinetics between the BPH14 LRR and $BISP^{125–241}$. BLI response profile of $BISP^{125–241}$ at different concentrations with the sensor-immobilized LRR. **e**, MST analysis of binding affinity between fluorescently labeled LRR, N14-LRR and BISP, $BISP^{26–124}$ and $BISP^{125–241}$. The K_d values show the binding between LRR and BISP, $BISP^{26–124}$. ΔF_{norm} is the normalized fluorescence minus the relative fluorescence of the unbound state. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **f**, MST analysis of the competitive binding between BISP and $BISP^{26–124}$ with LRR. ΔF_{norm} is the normalized fluorescence minus the relative fluorescence of the unbound state. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **g**, BISP protein levels in the independent *Bph14-Bisp* homozygous T_2 transgenic lines and *Bph14* control plants. Total proteins were immunoblotted and detected with anti-Myc and anti-ACTIN antibodies, respectively. **h**, Relative transcript levels of *Bph14*-activated SA biosynthesis- and signaling-related genes *OsICS1* and *OsNPR1*, and defense-related genes *OsPR1a*, *OsPR1b*, *OsPR5* and *OsPR10* in the *Bph14-Bisp* transgenic lines and WT *Bph14* plants. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **i**, Honeydew excretion by BPHs reared on the *Bph14-Bisp* transgenic lines and WT *Bph14* plants for 48 h. *P* values were derived by one-way ANOVA ($n = 20$, biologically independent plants). In box plots, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted. **j**, Survival rate of BPHs feeding on the *Bph14-Bisp* transgenic lines and WT *Bph14* plants for 9 days. *P* values were derived by one-way ANOVA. Data are mean ± s.e.m. ($n = 5$, biologically independent experiments). **k**, Two-host choice test of BPHs on the *Bph14-Bisp* transgenic plants and WT *Bph14* plants. *P* values were derived by one-way ANOVA. Data are mean ± s.e.m. ($n = 8$, biologically independent experiments). All the experiments were repeated two (**b**–**d**) or three (**a**,**g**,**i**–**k**) times, each giving similar results.

Source data

Extended Data Fig. 6 Agronomic traits of the *Bph14-Bisp* transgenic lines and WT *Bph14* plants.

a, Phenotypes of the *Bph14-Bisp* transgenic lines and *Bph14* plants at the heading stage. Scale bar, 30 cm. **b**, **c**, Plant height (**b**) and heading date (**c**) of the *Bph14-Bisp* transgenic lines and *Bph14* plants at the heading stage. *P* values were derived by one-way ANOVA. For **b**, $n = 18$, biologically independent samples. For **c**, $n = 30$, biologically independent samples. **d**–**j**, Agronomic traits of the *Bph14-Bisp* transgenic lines and WT *Bph14* plants at the mature stage. (**d**), Flag leaf length. (**e**), Flag leaf width. (**f**), The number of panicles per plant. (**g**), The number of filled grains per plant. (**h**), Percentage of filled grains. (**i**), 1,000-grain weight. (**j**), Grain yield per plant. One-way ANOVA ($n = 18$, biologically independent samples). In box plots in **b**–**j**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted.

Source data

Extended Data Fig. 7 The autophagic degradation of BISP was promoted by BPH14.

a, Dosage effect of BPH14 on BISP levels in rice protoplasts. BISP-MYC was expressed with or without BPH14-MYC in rice protoplasts. Total proteins were immunoblotted and detected with anti-Myc and anti-ACTIN antibodies, respectively. **b**, Immunoblots showing that proteasome degradation of WRKY72-HA was blocked by MG132 treatment. Total protoplast proteins were immunoblotted and detected with anti-HA and anti-ACTIN antibodies, respectively. **c**, Immunoblots showing that the autophagy degradation of Hd1-HA was blocked by treatments with autophagy inhibitors. Total protoplast proteins were immunoblotted and detected with anti-HA and anti-ACTIN antibodies, respectively. **d**, (*Upper*) Confocal images of CFP-ATG8f-labeled autophagic puncta structures in the absence (*left row*) or presence (*right row*) of ConA. Autophagosomes in cells co-expressing CFP-ATG8f with BISP, N14, or BPH14 (individually)

and CFP-ATG8f and BISP with N14 or BPH14 after agroinfiltration of *N. benthamiana* leaves are displayed. Red arrows indicate CFP-ATG8f-labeled autophagic puncta. (*Bottom*) Insets show enlarged autophagosome image at higher magnification. Scale bars, 25 μm. **e**, Average number of CFP-NbATG8-labeled autophagic puncta in the presence or absence ConA. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 6$, biologically independent experiments). **f**,**g**,**h**, Silencing of *NbATG6*, *NbPI3K* or *NbATG7* suppressed the production of CFP-ATG8f-labeled autophagic puncta. (**f**), Confocal images of CFP-ATG8f-labeled autophagic puncta structures co-localized with BISP-YFP. CFP-ATG8f, BISP-YFP and BPH14 were transiently co-expressed in the silenced leaves (VIGS-*NbATG6*, VIGS-*NbPI3K* and VIGS-*NbATG7*) or the non-silenced control leaves (VIGS-EV). Red arrows indicate CFP-ATG8f-labeled autophagic puncta co-localized with BISP-YFP. (*Bottom*) Insets show enlarged autophagosome image at higher magnification. Scale bars, 25 μm. (**g**), The transcript levels of *NbATG6*, *NbPI3K* and *NbATG7* by semi-quantitative RT-PCR. The transcript levels of *NbACTB* were used as a loading control. (**h**), Average number of CFP-NbATG8-labeled autophagic puncta in the silenced and control leaves. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 6$, biologically independent experiments). **i**, Immunoblot detection of OsATG8 in the non-infested *Bph14-Bisp* transgenic lines, *Bph14* and MH63 plants. Total proteins were immunoblotted and detected with anti-AtATG8A and anti-ACTIN antibodies, respectively. **j**, Relative transcript levels of *OsATG8s* (*OsATG8a, OsATG8b, OsATG8c*) in the non-infested *Bph14-Bisp* transgenic lines, *Bph14* and MH63 plants. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **k**, Relative transcript levels of *OsATG8s* (*OsATG8a, OsATG8b, OsATG8c*) in *Bph14* and MH63 plants after BPH feeding for different durations. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). **l**, Immunoblot detection of OsATG8 and OsNBR1 in the non-infested and BPH-infested MH63 and *Bph14* plants in the absence or presence of ConA treatment. Leaf sheaths total proteins were immunoblotted and detected with anti-AtATG8A, anti-AtNBR1 and anti-ACTIN antibodies, respectively. Numbers under the lanes (**a**,**c**,**i**,**l**) indicate protein abundance relative to that

of ACTIN, quantified by ImageJ. Experiments (**a**–**d**,**f**,**g**,**i**,**l**) were repeated three times, each giving similar results.

Source data

Extended Data Fig. 8 OsNBR1 played a role in the autophagic degradation of BISP by BPH14.

a, Immunoblot detection of proteins expressed in yeast using anti-HA and anti-Myc antibodies, respectively. Ponceau S staining was used as the loading control. Asterisks indicate specific bands detected by immunoblotting analysis. **b**, Confocal images of CFP-NbATG8f-labeled autophagic puncta structures in the absence (*left*) or presence (*right*) of ConA. The photos showing cells co-expressing CFP-NbATG8f and BISP with N14 and OsNBR1, BPH14, BPH14 and OsNBR1, or BPH14 and OsNBR1 mutant N1 after agroinfiltration of *N. benthamiana* leaves. Red arrows indicate CFP-NbATG8f-labeled autophagic puncta. (*Bottom Row*) Insets show enlarged autophagosome image at higher magnification. EV, Empty Vector. Scale bars, 25 µm. **c**, Average number of CFP-NbATG8-labeled autophagic puncta in the absence or presence of ConA. Data are mean ± s.d. ($n = 10$, biologically independent samples). Different letters indicate significant differences ($P < 0.05$, one-way ANOVA); exact P values are provided in Supplementary Table 4. **d**, Immunoblot detection of OsNBR1 levels in the non-infested *Bph14-Bisp* transgenic lines, *Bph14* and MH63 plants. Total proteins were immunoblotted and detected with anti-AtNBR1 and anti-ACTIN antibodies, respectively. Numbers under the lanes indicate protein abundance relative to that of ACTIN, quantified by ImageJ. **e**, Relative *OsNBR1* transcript levels in the non-infested *Bph14-Bisp* transgenic lines, *Bph14* and MH63 plants. *P* values were derived by one-way ANOVA. Data are means ± s.d. ($n = 3$, biologically independent experiments). **f**, Relative *OsNBR1* transcript levels in *Bph14* and MH63 plants after BPH infestation. *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). Experiments (**a**,**b**,**d**) were repeated at least three times, each giving similar results.

Source data

Extended Data Fig. 9 The PB1 domain, UBA domain and LIR motif of OsNBR1 played essential roles in BPH14-promoted autophagic degradation of BISP.

a, Y2H assay of the interactions between OsNBR1 and the CC, NB, and LRR domains of BPH14. DDO, SD/-Leu-Trp; TDO, SD/-Leu-Trp-His with 1 mM 3-AT. **b**, Co-IP assay of the interactions between OsNBR1 and the CC, NB, and LRR domains of BPH14 in rice protoplasts. OsNBR1-MYC and CC-HA, NB-HA or LRR-HA were co-expressed in rice protoplasts, and used in the co-IP assay with anti-Myc agarose; the input and immunoprecipitated proteins were detected with anti-HA, anti-Myc or anti-ACTIN antibodies. **c**, Schematic illustration of OsNBR1 and the N1-N7 mutant constructs used in this study. **d**, OsNBR1 mutants were unable to degrade BISP in rice protoplasts. BISP-MYC was co-expressed with BPH14-MYC and the full-length OsNBR1 (FL-HA) or its HA-tagged mutants (N1-N7-HA) in rice protoplasts. The levels of BISP-MYC protein were detected with anti-Myc antibody. **e**–**h**, Co-IP assays of the interactions between the OsNBR1 mutants and BISP (**e**), BPH14 (**f**), OsATG8c (**g**) and OsNBR1 (**h**) in rice protoplasts. BISP-MYC, BPH14-MYC, MYC-OsATG8c, or OsNBR1-MYC was co-expressed with OsNBR1(FL-HA) and its HA-tagged mutants (N1-N7-HA) in rice protoplasts. Proteins interacting with OsNBR1-HA or its mutants were identified in a co-IP assay with anti-MYC agarose; the input and immunoprecipitated proteins were detected with anti-HA, anti-Myc or anti-ACTIN antibodies. **i**, Summary of interactions between the OsNBR1 mutants and BISP, BPH14, OsATG8c, and OsNBR1, respectively. **j**, Verification of the mutations in the independent homozygous *OsNBR1* knockout (*Bph14*-O*snbr1*, *Osnbr1*$^{-}$) lines (*Bph14-Osnbr1*$^{-}$-1 and *Bph14-Osnbr1*$^{-}$-2) by PCR-based sequencing. **k**, Immunoblot detection of OsNBR1 proteins in *Bph14* and two *OsNBR1* knockout lines (*Bph14-Osnbr1*$^{-}$-1 and *Bph14-Osnbr1*$^{-}$-2). Total proteins were immunoblotted and detected with anti-AtNBR1 and anti-ACTIN antibodies, respectively. In **c**-**i**, FL, full-length of OsNBR1. N1-N7, the seven OsNBR1 mutants. Experiments were repeated three times (**a**,**b**,**d**–**h**,**k**) or two times (**j**) with similar results.

Extended Data Fig. 10 Effects of BPH infestation cessation in BPH14 activation and termination and the role OsNBR1 played in BPH resistance.

a, Relative transcript levels of *OsATG8s* and *OsNBR1* in *Bph14* and MH63 plants upon the cessation of BPH feeding. **b**, Relative transcript levels of *Bph14*-mediated resistance signaling genes, *OsICS1* and *OsNPR1*, in *Bph14* and MH63 plants upon the cessation of BPH feeding. **c**, Relative transcript levels of *Bph14*-mediated resistance signaling genes, *OsICS1* and *OsNPR1*, in the *Bph14-Osnbr1*$^{-}$ (*Bph14*-Os*nbr1*) lines (*Bph14-Osnbr1*$^{-}$-1 and *Bph14-Osnbr1*$^{-}$-2) and *Bph14* plants upon the cessation of BPH feeding. **d**, BPH weight gain on the *Bph14-Osnbr1*$^{-}$ lines (*Bph14-Osnbr1*$^{-}$-1 and *Bph14-Osnbr1*$^{-}$-2) and *Bph14* plants after 48 h of feeding. **e**, Honeydew excretion by BPHs reared on the *Bph14-Osnbr1*$^{-}$ lines and *Bph14* plants after 48 h of feeding. **f**, Two-host choice test of BPHs feeding on the *Bph14-Osnbr1*$^{-}$ lines and *Bph14* plants. **g**, Survival rates of BPHs feeding on the *Bph14-Osnbr1*$^{-}$ lines and *Bph14* plants for 8 days. In **a**,**b**,**c**, *P* values were derived by one-way ANOVA. Data are mean ± s.d. ($n = 3$, biologically independent experiments). In **d**,**e**, *P* values were derived by one-way ANOVA ($n = 18$, biologically independent samples). In **f**,**g**,*P* values were derived by one-way ANOVA. Data are mean ± s.e.m. ($n = 10$, biologically independent samples). In box plots in **d**,**e**, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. Individual data points are plotted.

Source data

Supplementary information

Supplementary Fig. 1

Uncropped blots and gel images.

Reporting Summary

Supplementary Fig. 2

The other two repeats of Fig. 1a–e,h–k.

Supplementary Table 1

New rice varieties introgressed with *Bph14* and authorized for release into production.

Supplementary Table 2

BPH14-interacting proteins from BPH identified in a Y2H assay.

Supplementary Table 3

List of primers used in this study.

Supplementary Table 4

Statistical summary—a summary of all statistical analysis is provided. Source data for all of the statistical analysis can be found in the source data files.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Fig. 5

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Guo, J., Wang, H., Guan, W. *et al.* A tripartite rheostat controls self-regulated host plant resistance to insects. *Nature* **618**, 799–807 (2023). https://doi.org/10.1038/s41586-023-06197-z

- Received: 06 March 2022
- Accepted: 11 May 2023
- Published: 14 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06197-z

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06197-z

- Article
- Open Access
- Published: 21 June 2023

Signalling by senescent melanocytes hyperactivates hair growth

- Xiaojie Wang ,
- Raul Ramos,
- Anne Q. Phan,
- Kosuke Yamaga,
- Jessica L. Flesher,
- Shan Jiang,
- Ji Won Oh,
- Suoqin Jin,
- Sohail Jahid,
- Chen-Hsiang Kuan,
- Truman Kt Nguyen,
- Heidi Y. Liang,
- Nitish Udupi Shettigar,
- Renzhi Hou,
- Kevin H. Tran,
- Andrew Nguyen,
- Kimberly N. Vu,
- Jennie L. Phung,
- Jonard P. Ingal,
- Katelyn M. Levitt,
- Xiaoling Cao,
- Yingzi Liu,
- Zhili Deng,
- Nobuhiko Taguchi,
- Vanessa M. Scarfone,
- Guangfang Wang,

- Kara Nicole Paolilli,
- Xiaoyang Wang,
- Christian F. Guerrero-Juarez,
- Ryan T. Davis,
- Elyse Noelani Greenberg,
- Rolando Ruiz-Vega,
- Priya Vasudeva,
- Rabi Murad,
- Lily Halida Putri Widyastuti,
- Hye-Lim Lee,
- Kevin J. McElwee,
- Alain-Pierre Gadeau,
- Devon A. Lawson,
- Bogi Andersen,
- Ali Mortazavi,
- Zhengquan Yu,
- Qing Nie,
- Takahiro Kunisada,
- Michael Karin,
- Jan Tuckermann,
- Jeffrey D. Esko,
- Anand K. Ganesan,
- Ji Li &
- …
- Maksim V. Plikus

Show authors

Nature **volume 618**, pages 808–817 (2023)

- 9628 Accesses

- 990 Altmetric

- Metrics details

Abstract

Niche signals maintain stem cells in a prolonged quiescence or transiently activate them for proper regeneration[1]. Altering balanced niche signalling can lead to regenerative disorders. Melanocytic skin nevi in human often display excessive hair growth, suggesting hair stem cell hyperactivity. Here, using genetic mouse models of nevi[2,3], we show that dermal clusters of senescent melanocytes drive epithelial hair stem cells to exit quiescence and change their transcriptome and composition, potently enhancing hair renewal. Nevus melanocytes activate a distinct secretome, enriched for signalling factors. Osteopontin, the leading nevus signalling factor, is both necessary and sufficient to induce hair growth. Injection of osteopontin or its genetic overexpression is sufficient to induce robust hair growth in mice, whereas germline and conditional deletions of either osteopontin or CD44, its cognate receptor on epithelial hair cells, rescue enhanced hair growth induced by dermal nevus melanocytes. Osteopontin is overexpressed in human hairy nevi, and it stimulates new growth of human hair follicles. Although broad accumulation of senescent cells, such as upon ageing or genotoxic stress, is detrimental for the regenerative capacity of tissue[4], we show that signalling by senescent cell clusters can potently enhance the activity of adjacent intact stem cells and stimulate tissue renewal. This finding identifies senescent cells and their secretome as an attractive therapeutic target in regenerative disorders.

Main

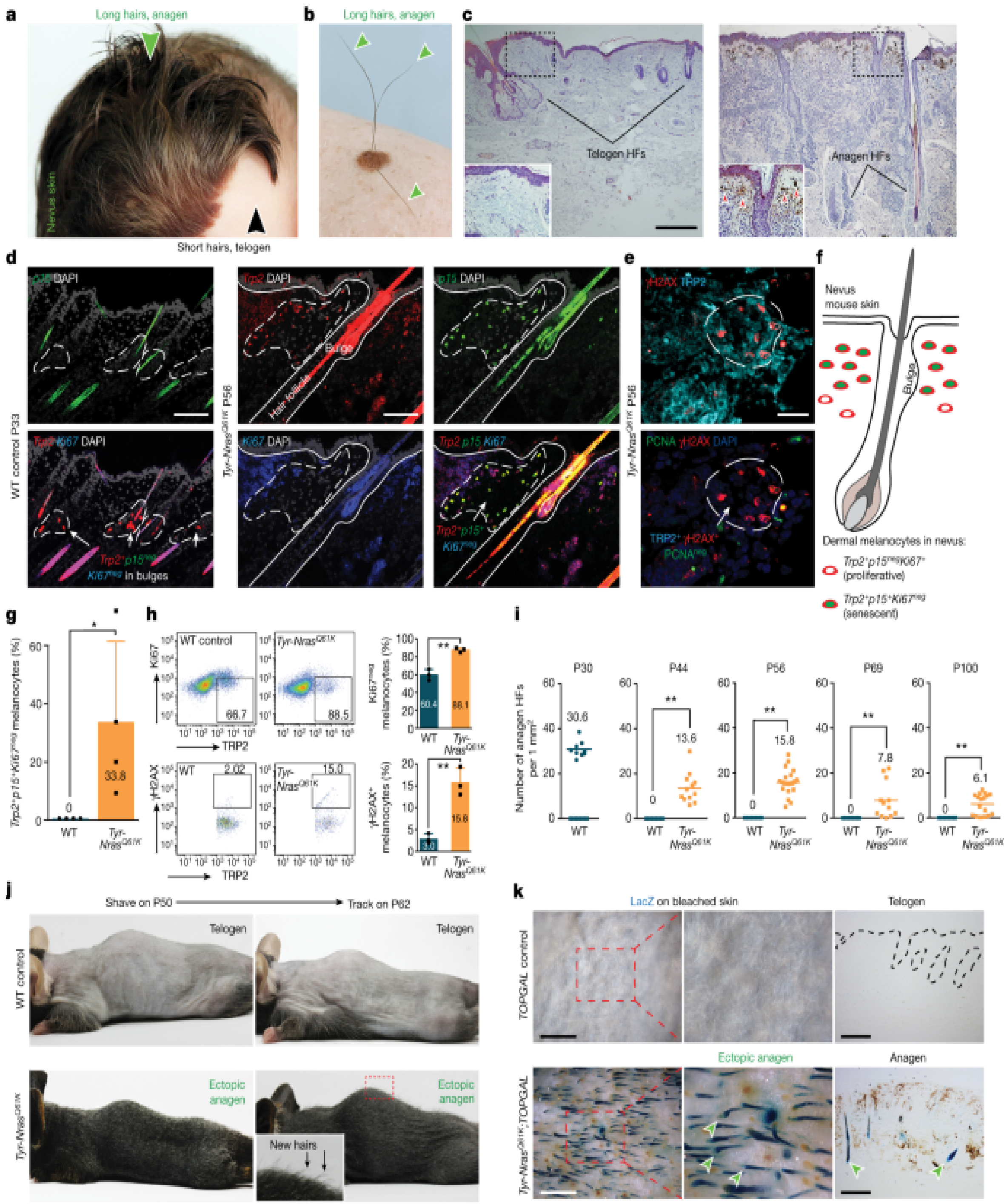
Stem cells (SCs) are critically required for long-term tissue maintenance and regeneration. To perform their function, SCs remain quiescent and transiently activate only when warranted, a switch that is tightly controlled. Immediate control is exerted by the short-range signalling niche[1]. In addition, activities of thousands of individual SC niches are coordinated by long-range signalling cues from the surrounding tissues[5]. Because long-range signals coordinate activities of many SC niches at once, any changes in them can profoundly alter the overall regenerative potential of an organ. However, which cell types can function as efficient long-range regulators of SCs is poorly understood.

Skin offers a valuable model system for studying these fundamental aspects of SC biology. Skin contains progenitor-rich hair follicles (HFs) that renew

in cycles[6]. Each cycle starts with SC activation[7] and requires signalling by the niche, featuring specialized dermal papilla fibroblasts[8]. Although in principle HFs are able to renew cyclically without external signalling inputs, many thousands of HFs physiologically coordinate their hair-making activities for the common goal of proper fur ‘manufacturing’[9]. Coordination is achieved via shared signalling between neighbouring HFs[10] and other non-hair skin cell types. The most prominent effects on hair renewal are exerted by skin adipocytes[11] and adipose progenitors[12]. This is possible because HFs and adipose tissue are close to each other and because they use some of the same signalling pathways—WNT, BMP, Hedgehog and PDGF—to regulate their cellular lineages. Innate and adaptive immune cells are also potent modifiers of hair growth dynamics[13,14].

Because cyclic hair renewal is tightly controlled at the level of SC quiescence, naturally occurring conditions of excessive hair growth are rare. Hairy pigmented nevi, both congenital (Fig. 1a,c) and acquired (Fig. 1b), are a type of benign skin lesion in humans that can show prominent hair growth. Despite being well known clinically, the mechanism behind excessive hair growth in nevi is not understood. Oncogene mutations, commonly in *Nras* (also known as *Alps4*) or *Braf*, in skin melanocytes induce nevi[15]. Mutant cells first transiently expand but subsequently activate oncogene-induced senescence (OIS)[16], giving rise to a spatially restricted lesion enriched for senescent cells. Once in full senescence, cells express a specialized secretome: the senescence-associated secretory phenotype (SASP)[17]. Several inflammatory cytokines and growth factors are part of the SASP, and their essential signalling roles are being rapidly recognized in normal embryonic development[18], cellular reprogramming[19], injury repair[20] and cancer progression[21,22]. We hypothesized that enhanced hair growth in hairy nevi is driven by activating signalling from dermal clusters of senescent melanocytes to HF SCs.

Fig. 1: Hyperactivation of hair growth in nevus skin.



a,**b**, Hair growth (arrowheads) is enhanced within congenital (7-month old; **a**) and acquired (42-year old; **b**) melanocytic nevi in humans. **c**, Facial HFs that commonly remain in telogen in normal skin (left) activate and enter new anagen in nevus skin (right). The red arrowheads mark dermal melanin. **d**,**f**,**g**, Compared with P33 WT anagen skin, P56 *Tyr-Nras*^{Q61K} skin

contained clusters of *Trp2*$^+$*p15*$^+$*Ki67*^{neg} melanocytes in the upper dermis. In **g,** *n* = 4; *P* = 0.0455668. **e**, P56 *Tyr-Nras*^{Q61K} skin contained clusters of TRP2$^+$γH2AX$^+$PCNA^{neg} dermal melanocytes. **h**, Compared with P30 WT anagen skin, P56 *Tyr-Nras*^{Q61K} skin showed significantly increased numbers of TRP2$^+$Ki67^{neg} (*n* = 3; *P* = 0.0019135) and TRP2$^+$γH2AX$^+$ melanocytes on cytometry (*n* = 3; *P* = 0.0028236). **i**–**k**, *Tyr-Nras*^{Q61K} mice displayed enhanced hair growth. At all postnatal time points examined (also see Extended Data Fig. 1), *Tyr-Nras*^{Q61K} skin contained many ectopic anagen HFs. Anagen HFs are quantified (**i**). In **i**, *n* = 9 at P30, *n* = 12 (*P* = 0.0000108) at P44, *n* = 21 (*P* = 0.0000000000183) at P56, *n* = 12 (*P* = 0.00329) at P69 and *n* = 17 (*P* = 0.0000239) at P100. In **j**, 12 days after shaving at P50, many new hairs grew in *Tyr-Nras*^{Q61K}, but not in WT mice. In **k**, at P56, *Tyr-Nras*^{Q61K};*TOPGAL* mice, but not control *TOPGAL* mice, showed many lacZ$^+$ anagen HFs (arrowheads). In **g**–**i**, *n* refers to biologically independent samples. Data are mean ± s.d. *P* values were calculated using unpaired one-tailed (**g**,**i**) or two-tailed (**h**) Student's *t*-test. **P* ≤ 0.05 and ***P* ≤ 0.01. Scale bars, 20 μm (**e**), 100 μm (**d**), 500 μm (**c**), 1 mm (wholemount; **k**) and 200 μm (histology; **k**). The image in part **a** is reproduced with permission from S. Liber.

Source data

Senescent cells activate hair growth

First, we asked whether mouse models for melanocytic nevi replicate enhanced hair growth. We studied two established models: constitutive *Tyr-Nras*^{Q61K} mice[2], which model congenital nevi, and inducible *Tyr-CreER*^{T2};*Braf*^{V600E} mice[3], which model acquired nevi. In both models, oncogenes are overexpressed from the *Tyr* enhancer–promoter regulatory region that is highly specific to neural crest-derived melanocytes. Normal hair growth in mice is coordinated: large groups of HFs jointly transition from the resting phase (telogen) to the active growth phase (anagen) and then via the regression phase (catagen) back into telogen[9,11]. This coordination causes HF SCs to spend a large portion of their lifecycle in quiescence, only transiently activated to regenerate new hairs within discrete HF groups. Resting HFs house melanocyte SCs, located in the shared niche

with epithelial SCs, whereas growing HFs also contain activated, pigment-producing melanocytes at their base. *Tyr-Nras*^{Q61K} mice, whose dermis but not HFs themselves become populated by senescent melanocytes identified as non-proliferative *p15*$^{+}$ (Fig. 1d,f,g) and non-proliferative γH2AX^{+} melanocytes (Fig. 1e,h), showed dramatically accelerated hair growth, with many ectopic anagen HFs present at any given time ($n = 3$ per time point) (Fig. 1i–k and Extended Data Fig. 1). In control mice, dorsal HFs were in first anagen at postnatal day 15 (P15) (Extended Data Fig. 1a), first telogen by P23 (Extended Data Fig. 1b) and second anagen by P36 (Extended Data Fig. 1c). After that, HFs entered a lengthy second telogen spanning P44–P69 (Extended Data Fig. 1d–g). By contrast, at all time points examined, *Tyr-Nras*^{Q61K} skin contained ectopic anagen HFs (Fig. 1i,j and Extended Data Fig. 1), which were numerous even at P100 (Extended Data Fig. 1h). The ectopic anagen phenotype was especially visible in *Tyr-Nras*^{Q61K};*TOPGAL* mice ($n = 4$), where all anagen HFs strongly activated the *TOPGAL* WNT reporter and stained positive for lacZ (Fig. 1k). Ectopic anagen HF density in *Tyr-Nras*^{Q61K} mice varied between the time points, but on average it was 35.4% relative to synchronous anagen HF density in P30 wild-type (WT) skin (Fig. 1i). We crossed *Tyr-Nras*^{Q61K} mice onto an albino *Tyr*(*C-2J*) background carrying a mutation in the *Tyr* gene. Despite the lack of melanin, albino *Tyr-Nras*^{Q61K} mice displayed ectopic anagen at both P56 and P100 (Extended Data Fig. 1j), indicating that it is not excessive melanogenesis but rather senescent melanocytes that are necessary for the nevus hair phenotype.

Next, we modelled early acquired nevi in *Tyr-CreER*^{T2};*Braf*^{V600E} mice that were treated with tamoxifen either early at P2–P4 or late at P21–P25. Unlike induced control animals, induced mutant mice accumulated clusters of senescent non-proliferative *p15*$^{+}$ (Extended Data Fig. 2a), non-proliferative γH2AX^{+} (Extended Data Fig. 2b,c) and non-proliferative *p16*$^{+}$ (Extended Data Fig. 2d) melanocytes in the dermis adjacent to HFs. Mutant mice induced at P2–P4 displayed prominent ectopic anagen at P44, P56, P69 and P100 ($n = 4$ per time point) (Extended Data Figs. 1k and 4a,b). Across time points, they averaged 35.7% anagen HFs relative to P30 WT skin, which closely phenocopied congenital *Tyr-Nras*^{Q61K} mutants. Likewise, mutant mice treated with tamoxifen at P21–P25 also showed prominent ectopic anagen starting at P56 ($n \geq 3$ per time point) (Extended Data Fig. 4c,d). We

更多杂志请加vx
几百种杂志持续更新

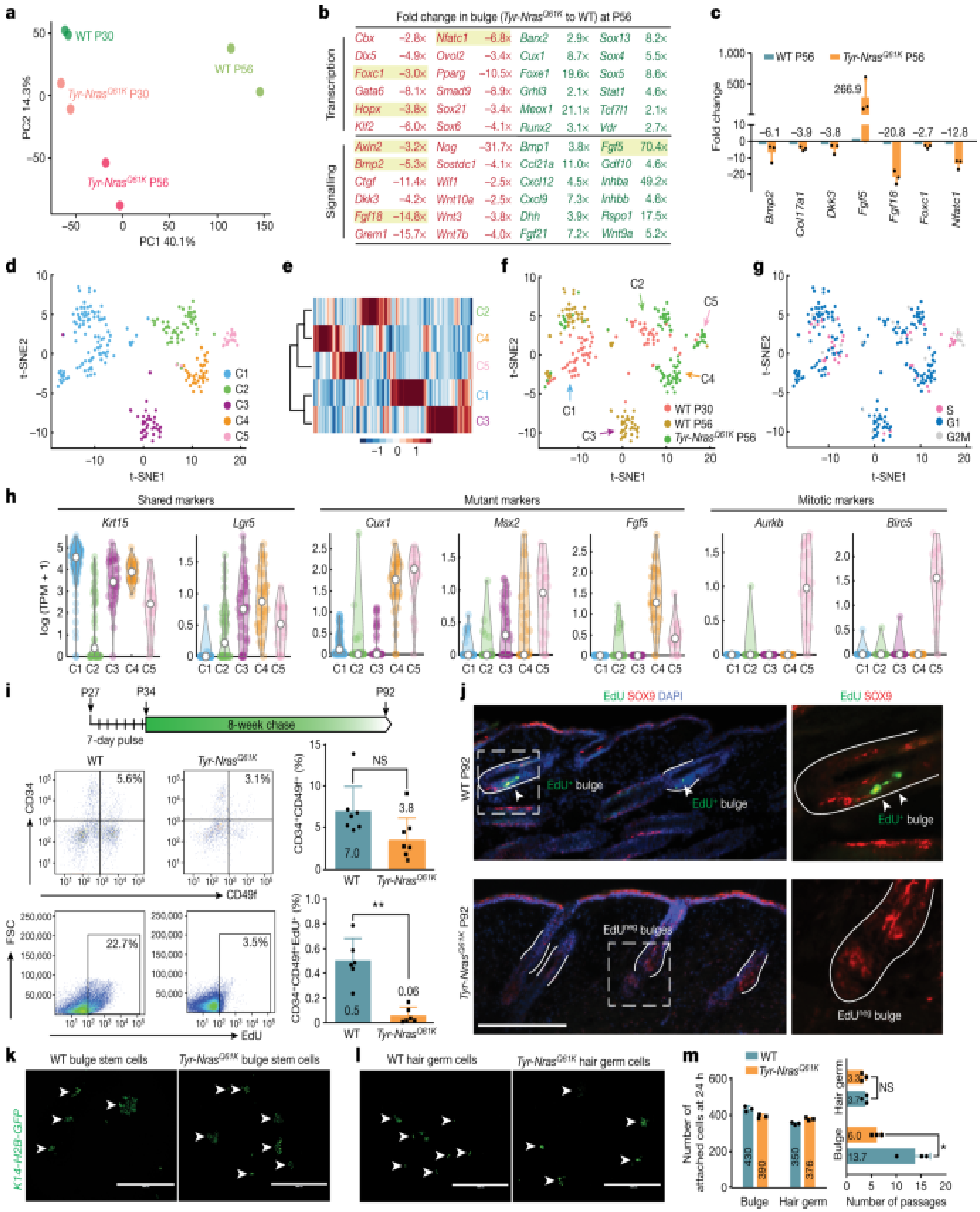
also asked whether injection of nevus-derived melanocytes into normal telogen skin would be sufficient to induce ectopic anagen. We sorted tdTomato^{+} melanocyte lineage cells from the skin of congenital *Tyr-Nras*^{Q61K};*Tyr-CreER*^{T2};*tdTomato* (Extended Data Fig. 3a) and acquired *Tyr-CreER*^{T2};*Braf*^{V600E};*tdTomato* mice (Extended Data Fig. 3c). Intradermal injection of sorted cells from both nevus mouse models into telogen skin of *SCID* mice ($n = 4$ each) induced new anagen within 21 days (Extended Data Fig. 3b,e), albeit their continued senescent status at the grafted site was not verified. Yet, by contrast, injection of sorted cells from control *Tyr-CreER*^{T2};*tdTomato* mice isolated during both telogen (P56) and anagen (P33) did not activate new anagen in *SCID* host skin ($n = 4$ each) (Extended Data Fig. 3d,f–i). We also generated senescent β-galactosidase-positive (β-Gal^{+}) melanocytes by exposing primary CD117^{+} newborn mouse melanocytes to H_2O_2 in vitro (Extended Data Fig. 3j–l). Unlike control cultured melanocytes ($n = 7$), DiI-labelled H_2O_2-treated melanocytes induced new anagen in telogen *SCID* skin 21 days after injection ($n = 6$) (Extended Data Fig. 3m–o). We also subcutaneously treated mice with the small-molecule BCL-2 inhibitor ABT-737, which in P56 *Tyr-Nras*^{Q61K} mice, induced prominent apoptosis of melanocytes but did not affect the abundance of HF SCs ($n = 5$) (Extended Data Fig. 5c,d), and in P33 WT mice did not delay normal anagen timing ($n = 7$) (Extended Data Fig. 5e). By contrast, ABT-737 treatment of *Tyr-Nras*^{Q61K} mice significantly reduced ectopic anagen HFs at P56 ($n = 6$) (Extended Data Fig. 5a,b), which we attribute to nevus melanocyte depletion. Next, we studied *K14-Edn3* and *K14-Kitl* mice, which, respectively, showed expansion in dermal and epidermal melanocytes that is not driven by oncogene mutation. Both mouse models showed normal hair cycle progression, with synchronized anagen at P36 ($n = 3$ per model) and synchronized telogen at P56 ($n = 3$ per model) (Extended Data Fig. 5f–i). Last, we induced *Trp53* (also known as *p53*) deletion in melanocytes, which despite being an oncogenic stimulation, did not induce OIS[17], unlike *Nras*^{Q61K} or *Braf*^{V600E} overexpression. Analogous to control mice, HFs in tamoxifen-treated P56 *Tyr-CreER*^{T2};*Trp53*$^{fl/fl}$ mice remained in telogen ($n = 3$) (Extended Data Fig. 5j). Together, our data show that congenital and acquired mouse models for melanocyte OIS reproduce the enhanced hair growth that is clinically observed in human hairy pigmented nevi and that senescent dermal

melanocytes, but not normal melanocytes, are necessary and sufficient to hyperactivate HF renewal.

Senescence disrupts SC quiescence

We next asked how bona fide HF bulge SCs are affected by the nevus environment. We profiled their transcriptomes by RNA sequencing (RNA-seq) at P30 and P56, when WT HFs are in anagen and telogen, respectively. Bulge SCs were isolated as $GFP^{+}CD34^{+}Pcad^{low}$ cells both from *K14-H2B-GFP* control mice and *Tyr-Nras*^{Q61K};*K14-H2B-GFP* mutant mice, in which CD34 and Pcad maintain WT expression patterns (Extended Data Fig. 6a,b). RNA-seq revealed prominent gene expression differences between *Tyr-Nras*^{Q61K} and control bulge SCs (Fig. 2a, Extended Data Fig. 6c and Supplementary Table 1). The largest differences were seen at P56, with mutant SCs downregulating and upregulating 973 and 1,159 genes, respectively. Depleted gene ontology categories for mutant SCs included cell cycle block, circadian rhythm, and WNT and JAK–STAT suppression, whereas enriched categories contained cell cycle, cell migration, WNT signalling and skin development (Extended Data Fig. 6d and Supplementary Table 1). These gene ontology signatures indicate that *Tyr-Nras*^{Q61K} bulge SCs lose quiescence. At the gene level, multiple quiescence markers, including *Axin2*, *Bmp2*, *Col17a1*, *Ctgf*, *Fgf18*, *Foxc1*, *Grem1*, *Nfatc1* and *Wif1*, were downregulated in P56 *Tyr-Nras*^{Q61K} SCs (Fig. 2b,c and Supplementary Table 1).

Fig. 2: Hair SCs within nevus skin lose quiescence.



a, On RNA-seq analysis, *Tyr-Nras*^{Q61K} bulge SCs differ from P30 and P56 WT bulge SCs. A principal component analysis plot is shown. See Extended Data Fig. 6. **b**, A list of selected downregulated (red) and upregulated (green) genes at P56 and *Tyr-Nras*^{Q61K} to WT fold change values. **c**, qRT–

PCR of selected differentially expressed genes from **a**. $n = 3$. **d**, t-Distributed stochastic neighbour embedding (t-SNE) analysis on single-cell RNA-seq data for P30 and P56 WT and P56 *Tyr-Nras*^{Q61K} bulge SCs. Cells form five clusters: C1 to C5. **e**, Cladogram showing relative cluster similarity. **f**, t-SNE plot colour-coded by sample source. **g**, t-SNE plot colour-coded by inferred cell cycle state. **h**, Violin plots for selected genes. See Extended Data Fig. 6. TPM, transcripts per million. **i**, EdU pulse-chase analysis on bulge SCs. Unlike total numbers of $CD34^+CD49f^+$ bulge SCs (top), their EdU^+ label-retaining subset reduced significantly in *Tyr-Nras*^{Q61K} versus control mice (bottom). $n = 7$ ($P = 0.061857$) for $CD34^+CD49f^+$ SCs and $n = 6$ ($P = 0.0002048$) for $CD34^+CD49f^+EdU^+$ SCs. See Extended Data Fig. 7. FSC, forward scatter. **j**, Unlike WT, *Tyr-Nras*^{Q61K} HFs from part **i** lacked EdU^+SOX9^+ bulge SCs (yellow). **k**,**l**, Attachment rates for the *K14-H2B-GFP*$^+$ bulge (**k**) and hair germ (**l**) cells were compatible between WT and *Tyr-Nras*^{Q61K} mice. Arrowheads mark cell colonies. **m**, Compared with WT, *Tyr-Nras*^{Q61K} bulge SCs prominently reduced serial passaging potential, whereas it was unaltered for hair germ progenitor cells. $n = 3$ ($P = 0.5185185$) for hair germ cells and $n = 3$ ($P = 0.0168963$) for bulge cells. In **c**,**i**,**m**, n refers to independent experiments. P values were calculated using unpaired two-tailed Student's t-test. Not significant (NS), $P \geq 0.05$, $^*P \leq 0.05$ and $^{**}P \leq 0.01$. Scale bars, 100 µm (**j**) and 1 mm (**k**,**l**).

Source data

To confirm that the *Tyr-Nras*^{Q61K} bulk RNA-seq signature is not being simply dominated by near-normal activated SCs from ectopic anagen HFs, we compared P56 mutant with P30 anagen and P56 telogen WT bulge cells by single-cell RNA-seq. WT cells from P30 and P56 formed the shared cluster C1 and two phase-specific clusters: anagen-specific C2 and telogen-specific C3 (Fig. 2d–f). Upon marker analysis, C1 cells matched the signature of inner bulge cells, which includes *Chit1*, *Krt6a* and *Krt80*, whereas both C2 and C3 cells matched that of outer bulge bona fide SCs, which includes *Col18a1*, *Krt17*, *Lhx2*, *Tcf7l2* and *Vdr*[23] (Fig. 2h, Extended Data Fig. 6e–j and Supplementary Tables 2 and 3). P56 mutant bulge cells dramatically altered their composition relative to WT cells; some cells contributed to the shared inner bulge cluster C1, others to the WT anagen-specific outer bulge cluster C2, whereas many cells formed two new mutant-

specific clusters C4 and C5, which retained a core outer bulge signature (Fig. 2f and Extended Data Fig. 6e). No mutant cells contributed to the WT telogen-specific outer bulge cluster C3, which has a quiescent gene expression signature, including *Bmp2*, *Col17a1*, *Ctgf*, *Grem1*, *Nfatc1*, *Tgm5* and *Wif1* (Extended Data Fig. 6f). Loss of quiescence by mutant-specific outer bulge SCs was further evident from inferred cell cycle analysis: C5 cells were exclusively in S and G2/M phases (Fig. 2g) with prominently upregulated mitotic markers (Fig. 2h and Supplementary Table 3). Given that *Tyr-Nras*^{Q61K} skin contains a mixture of anagen and telogen HFs, the disappearance of WT telogen-specific C3 outer bulge cells supports the loss of quiescence by mutant telogen SCs. Outer bulge marker similarities between clusters C2 to C5 suggest that in the presence of nevus melanocytes, normally quiescent telogen SCs transition to a uniquely activated state.

Next, we confirmed loss of quiescence in functional assays. For pulse and pulse-chase experiments, which measure the cell cycle status of cells, mice were treated with EdU between P27 and P34, when WT HFs are in anagen and their SCs proliferate. Four hours after the EdU pulse, *Tyr-Nras*^{Q61K} mice displayed bulge SC labelling efficiency that was compatible with WT SCs (Extended Data Fig. 7a,b). However, in a pulse-chase assay, there was a prominent loss of EdU-retaining SCs in *Tyr-Nras*^{Q61K} mice as noted upon analysis at P92 ($n = 4$ per genotype) (Fig. 2i,j). We then performed a clonogenic assay, which measures long-term proliferative potential by cultured cells and identifies SCs on the basis of them being able to form large clones over many serial passages. We show that the attachment ability of *Tyr-Nras*^{Q61K} bulge SCs was similar to that of WT SCs, but their serial passaging potential was compromised; mutant SCs supported 6 passages ($n = 3$) compared with 13.7 passages for WT SCs ($n = 3$) (Fig. 2k,m). A decrease in passaging potential by bulge SCs indicates their faster proliferative exhaustion, a likely consequence of their long-term hyperproliferative status in vivo before culture. Attachment rates and passaging potential, however, did not differ between mutant ($n = 3$) and WT mice for hair germ cells, a short-lasting population of epithelial progenitors in telogen HFs ($n = 3$) (Fig. 2l,m).

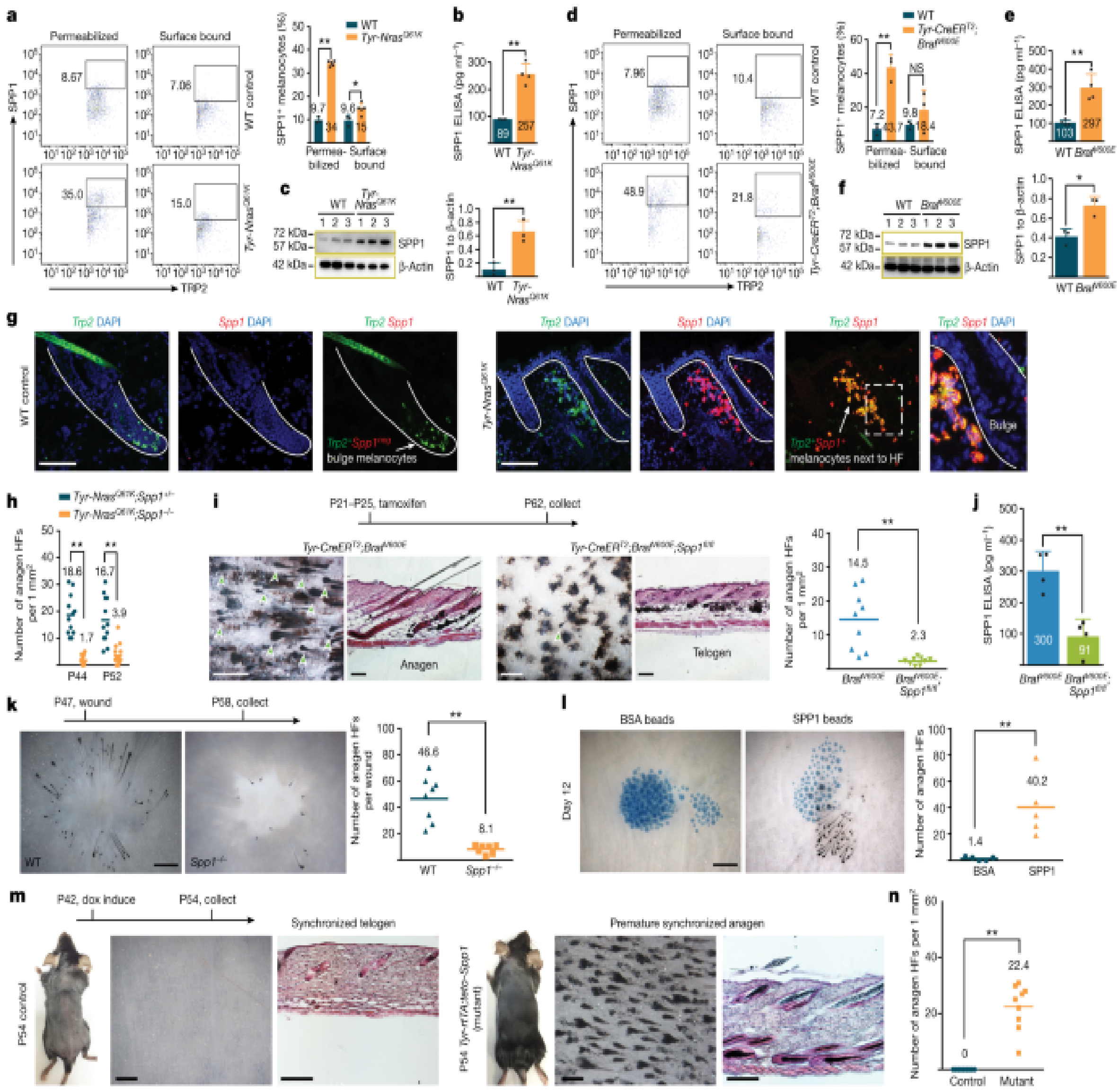
Osteopontin stimulates hair growth

Next, we asked which signalling factors are expressed by nevus melanocytes. We isolated the melanocyte lineage as $tdTomato^{+}$ cells from the tamoxifen-induced *Tyr-Nras*^{Q61K};*Tyr-CreER*^{T2};*tdTomato* mutant and *Tyr-CreER*^{T2};*tdTomato* control skin. P56 mutant cells were compared with both P30 anagen and P56 telogen WT cells on bulk RNA-seq (Extended Data Fig. 7c,d). This strategy identified 598 mutant-specific upregulated genes, and also excluded genes regulated as part of the normal hair cycle. Mutant-specific genes were enriched for gene ontology terms, including ageing, WNT suppression, cell cycle block and mitotic division (Extended Data Fig. 7f and Supplementary Table 4). Consistent with dermal clusters of mutant melanocytes undergoing OIS, they upregulated tumour suppressors *Cdkn2b* (also known as *p15*), *Lzts1*, as well as *Cdkn3*, *H2afx* and the mitosis-associated genes *Aurka/b*, *Cdca3/8*, *Cdc20/25c*, *Cenpa*, *Mad2l1*, *Ncaph*, *Knstrn*, *Plk1*, *Psrc1* and *Reep4* (Extended Data Fig. 7g,i). Upregulation of mitosis-associated genes is consistent with the fact that oncogene-stimulated melanocytes enter OIS via a mitotic arrest pathway, rather that via G0 phase[24]. Focusing on the secretome, we identified 27 signalling factors specifically upregulated in nevus melanocytes, including the BMP members *Bmp4* and *Fstl1*, the WNT members *Frzb*, *Wif1* and *Wisp1*, the IGF regulators *Igfbp2/4/7*, as well as *Dhh*, *Fgf7*, *Spp1* (also known as osteopontin) and *Tnf* (Extended Data Fig. 7e). Of note, 68% of the secretome genes enriched in *BRAF*^{V600E}-induced human senescent melanocytes in vitro[2] and 71% of the core in vitro SASP factors[17] were represented in the transcriptome of P56 *Tyr-Nras*^{Q61K} melanocytes (Extended Data Fig. 7h).

Spp1 was one of the topmost upregulated signalling transcripts in nevus melanocytes on RNA-seq. We confirmed this change at the protein level in sorted melanocytes from both the congenital and the acquired nevus mouse models. On cytometry, SPP1 levels were significantly increased in melanocytes from P56 *Tyr-Nras*^{Q61K} mice ($n = 3$) (Fig. 3a) and from tamoxifen-induced *Tyr-CreER*^{T2};*Braf*^{V600E} mice relative to control melanocytes at five time points between P44 and P100 ($n = 3$ each) (Fig. 3d and Extended Data Fig. 4f–i). Significantly increased SPP1 levels in P56 *Tyr-Nras*^{Q61K} and in P69 *Tyr-CreER*^{T2};*Braf*^{V600E} melanocytes were confirmed by western blot ($n = 3$ each) (Fig. 3c,f). Significant increase in SPP1 secretion was observed by ELISA on day 5 cultures of primary

melanocytes sorted from P56 *Tyr-Nras*^{Q61K} mice ($n = 3$) (Fig. 3b) and from *Tyr-CreER*^{T2};*Braf*^{V600E} mice at four time points between P56 and P100 relative to control melanocyte cultures ($n = 3$ each) (Fig. 3e and Extended Data Fig. 4e). On staining, clusters of *Trp2*$^{+}$*Spp1*$^{+}$ melanocytes were observed in the upper dermis adjacent to bulge regions of HFs only in nevus mice, both congenital (Fig. 3g) and acquired (Extended Data Fig. 4l), but not in control mice. Consistent with published gene expression analyses[25], lacZ staining in *Spp1*$^{+/-}$ mice (which carry β-Gal knock-in) shows that *Spp1* expression in normal skin at homeostasis is very restricted, largely limited to dermal papilla fibroblasts of HFs (Extended Data Fig. 8a–c). Together, the above data support that SPP1 is an upregulated signalling factor in dermal clusters of nevus melanocytes.

Fig. 3: Secretome of nevus melanocytes contains SPP1 that promotes hair growth.



a,**d**, On cytometry, SPP1 was increased in P56 *Tyr-Nras*^{Q61K} (**a**) and P69 *Tyr-CreER*^{T2};*Braf*^{V600E} (**d**) melanocytes. In **a**, for the permeabilized condition, $n = 3$ in WT and $n = 5$ in *Tyr-Nras*^{Q61K} ($P = 0.000000115$); for the surface-bound condition, $n = 3$ in WT and $n = 5$ in *Tyr-Nras*^{Q61K} ($P = 0.0257$). In **d**, for the permeabilized condition, $n = 3$ ($P = 0.001397$); for the surface-bound condition, $n = 3$ ($P = 0.2888$). See Extended Data Fig. 4. **c**,**f**, On western blot, SPP1 levels were increased in P56 *Tyr-Nras*^{Q61K} (**c**) and P69 *Tyr-CreER*^{T2};*Braf*^{V600E} (**f**) melanocytes. In **c**, $n = 3$; $P = 0.00784$. In **f**, $n = 3$; $P = 0.0109$. Uncropped gels are shown in Supplementary Fig. 1. **b**,**e**, On ELISA, SPP1 levels increased in day 5 cultures of P56 *Tyr-Nras*^{Q61K} (**b**) and P69 *Tyr-CreER*^{T2};*Braf*^{V600E} (**e**) melanocytes. In **b**, $n = 3$ in WT and $n = 4$ in

Tyr-Nras^{Q61K}; $P = 0.00072$. In **e**, $n = 4$; $P = 0.00224$. See Extended Data Fig. 4e. **g**, Unlike WT, *Tyr-Nras*^{Q61K} skin contained *Trp2*$^{+}$*Spp1*$^{+}$ melanocytes adjacent to HF bulges. **h**, Anagen HF quantification in *Tyr-Nras*^{Q61K};*Spp1*$^{-/-}$ versus *Tyr-Nras*^{Q61K};*Spp1*$^{+/-}$ control mice. At P44, $n = 12$ in control and $n = 14$ in *Tyr-Nras*^{Q61K};*Spp1*$^{-/-}$ ($P = 0.0000000191$); at P56, $n = 12$ in control and $n = 15$ in *Tyr-Nras*^{Q61K};*Spp1*$^{-/-}$ ($P = 0.0000195$). **i**, *Tyr-CreER*^{T2};*Braf*^{V600E};*Spp1*$^{fl/fl}$ mice showed hair cycle quiescence rescue. Representative samples (left) and quantification (right) are displayed. $n = 9$; $P = 0.000731$. Arrowheads mark anagen HFs. **j**, On ELISA, SPP1 levels were reduced in day 5 cultures of *Tyr-CreER*^{T2};*Braf*^{V600E};*Spp1*$^{fl/fl}$ versus *Tyr-CreER*^{T2};*Braf*^{V600E} melanocytes. $n = 4$; $P = 0.00242$. **k**, *Spp1*$^{-/-}$ mice showed reduced wound-induced hair growth. Representative samples (left) and quantification (right) are displayed. $n = 8$ in WT and $n = 7$ in *Spp1*$^{-/-}$; $P = 0.0000575$. **l**, Unlike BSA-soaked beads (blue), SPP1-soaked beads induced anagen in WT skin 12 days after injection. Representative samples (left) and quantification (right) are displayed. $n = 5$; $P = 0.00562$. **m**,**n**, Unlike control, doxycycline (dox)-treated P54 *Tyr-rtTA*;*tetO-Spp1* mice displayed premature anagen. Representative mice (**m**) and quantification (**n**) are displayed. In **n**, $n = 9$; $P = 0.000000377$. In **b**,**c**,**e**,**f**,**j**, *n* refers to independent experiments. In **a**,**d**,**h**,**i**,**k**,**l**,**n**, *n* refers to biologically independent samples. Data are mean ± s.d. *P* values were calculated using unpaired two-tailed Student's *t*-test. NS, $P \geq 0.05$, $^{*}P \leq 0.05$ and $^{**}P \leq 0.01$. Scale bars, 100 µm (**g**), 200 µm (histology; **i**,**m**) and 500 µm (wholemount; **i**,**k**,**l**,**m**).

Source data

Next, we asked whether SPP1 has a functional role in hairy nevus phenotype and whether it is sufficient to induce new hair growth. Using *Tyr-Nras*^{Q61K};*Spp1*$^{-/-}$ mice, we showed that a germline loss-of-function mutation in *Spp1* is sufficient to rescue hair cycle quiescence in congenital nevus skin. Compared with *Tyr-Nras*^{Q61K} mice, whose HFs start cycling ectopically already at P23 (Extended Data Fig. 1b), ectopic anagen in *Tyr-Nras*^{Q61K};*Spp1*$^{-/-}$ mice is largely prevented ($n = 6$ per time point) (Fig. 3h and Extended Data Fig. 8d–g). We also generated *Tyr-CreER*^{T2};*Braf*^{V600E};*Spp1*$^{fl/fl}$ mice, in which tamoxifen treatment induces a conditional *Spp1* loss-of-function mutation in melanocytes along with

oncogenic BRAF stimulation. We showed that melanocyte-specific *Spp1* deletion largely prevented ectopic hair cycle in P62 *Spp1*-deficient nevus mice compared with *Spp1*-intact nevus control animals ($n = 5$ each) (Fig. 3i), and that this correlated with a significant, approximately 70%, decrease in SPP1 secretion in primary melanocyte culture by ELISA ($n = 3$) (Fig. 3j). Partial SPP1 loss is attributed to incomplete efficiency of *CreER*-based recombination.

Unlike at homeostasis, SPP1 becomes prominently upregulated in skin wounds, both in wound fibroblasts[26] and wound macrophages[27]. Considering this, we asked whether it mediates wound-induced hair growth phenomenon, when HFs at the wound margin enter premature anagen. Indeed, compared with WT mice ($n = 8$), *Spp1*$^{-/-}$ mice showed significantly fewer anagen HFs at the margin of 5-mm wounds 11 days post-wounding ($n = 7$) (Fig. 3k). Ectopic anagen was prominently induced 12 days after intradermal injection of SPP1-soaked beads in WT mice compared with BSA-soaked control beads ($n = 5$ each) (Fig. 3l). Moreover, premature anagen was activated by P54 in *Tyr-rtTA*;*tetO-Spp1* mice, induced with doxycycline starting from P42. Compared with doxycycline-treated control mice, which remained in telogen, *Tyr-rtTA*;*tetO-Spp1* mice displayed broad anagen activation ($n = 3$ mice each) (Fig. 3m,n). Therefore, SPP1 is sufficient to induce new hair growth and it mediates hair growth activation in at least two skin states: melanocytic nevus and wound healing.

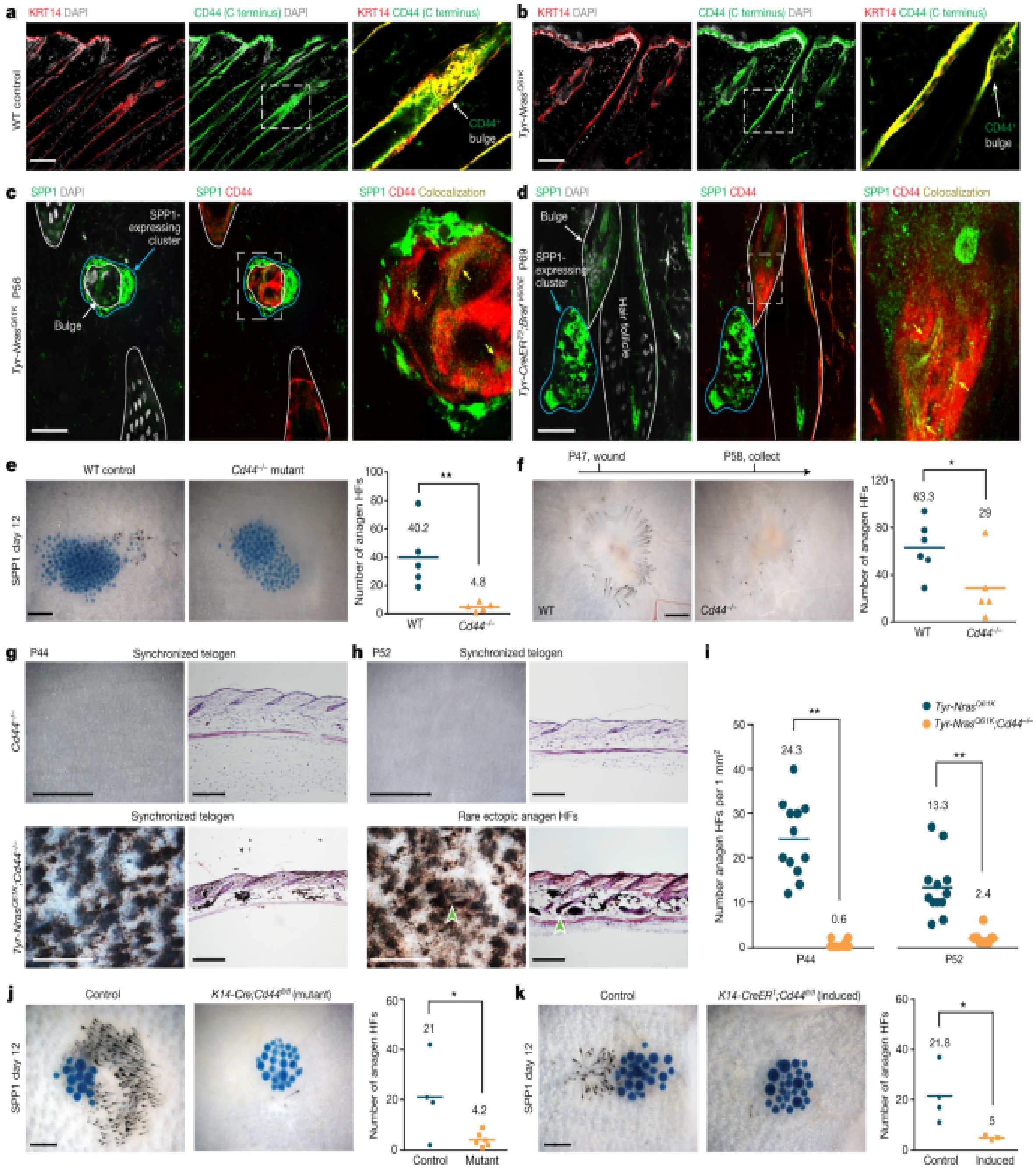
CD44 mediates the osteopontin effect

SPP1 signals via distinct binding sites to its cognate receptors: β-integrins and CD44 (also known as CSPG8). Of these, CD44 is an established stemness marker in several cancer types, where it promotes proliferation, invasiveness and radio-resistance[28]. SPP1 preferentially binds to alternatively spliced CD44v isoforms, which show enrichment in bulge SCs on RNA-seq (Extended Data Fig. 9a). In response to SPP1, CD44 undergoes proteolytic cleavage by γ-secretase, which releases its nuclear-targeted intracellular domain (CD44-ICD), thus coactivating HIF1A, EPAS1, EP300 and CREBBP to regulate gene expression[28]. *Mmp9*, a direct downstream target of CD44-ICD signalling[29], is one of the top upregulated genes in *Tyr-Nras*^{Q61K} bulge SCs (Extended Data Figs. 6c and 9b) and bulge SCs retain

high expression of all γ-secretase subunits as well as CD44-ICD-binding transcriptional factors (Extended Data Fig. 9c,d).

We asked whether CD44 mediates hair growth hyperactivation in the nevus. Consistent with previous single-cell RNA-seq profiling, CD44 is prominently expressed across all epithelial compartments of the HF[27], including in bulge SCs, both in control and *Tyr-Nras*^{Q61K} mice (Fig. 4a and Extended Data Fig. 9e). At the protein level, SPP1 colocalizes with CD44 in bulge SCs in both *Tyr-Nras*^{Q61K} and *Tyr-CreER*^{T2};*Braf*^{V600E} mice (Fig. 4c,d). Next, we asked whether *Cd44* deletion compromises bulge SC abundance and proliferative potential. The percentage of either total $CD34^{+}CD49f^{+}$ bulge SCs or their EdU-labelled subset after 7 days of EdU pulse did not significantly change in germline *Cd44*$^{-/-}$ mutant versus control mice ($n = 3$ each) (Extended Data Fig. 9f,g) as well as in epithelial-specific constitutive *K14-Cre*;*Cd44*$^{fl/fl}$ mutant versus control mice ($n = 3$ each) (Extended Data Fig. 9i,j). Also unchanged was the in vitro colony-forming potential by sorted bulge SCs both from *Cd44*$^{-/-}$ and *K14-Cre*;*Cd44*$^{fl/fl}$ mice versus control animals ($n = 6$ each) (Extended Data Fig. 9h,k). Therefore, loss of CD44 alone does not compromise key bulge SC properties. Next, we asked whether CD44 function is required for HF response to SPP1. Indeed, anagen induction in response to SPP1-soaked beads was significantly suppressed in *Cd44*$^{-/-}$ versus control mice ($n = 5$ each) (Fig. 4e). Likewise, significantly fewer anagen HFs were induced at the wound margin of *Cd44*$^{-/-}$ mutant ($n = 5$) versus control mice ($n = 6$) (Fig. 4f). Furthermore, *Cd44* deletion in *Tyr-Nras*^{Q61K};*Cd44*$^{-/-}$ mice led to rescue of ectopic hair cycling, phenocopying the effect of *Spp1* deletion in the *Tyr-Nras*^{Q61K} background (Fig. 4g–i and Extended Data Fig. 9l,m). Loss of SPP1 responsiveness in the soaked bead experiment was also phenocopied upon epithelial-specific *Cd44* deletion in *K14-Cre*;*Cd44*$^{fl/fl}$ as well as in tamoxifen-inducible *K14-CreER*^{T};*Cd44*$^{fl/fl}$ mice. Compared with SPP1-treated control mice ($n = 4$ each), the numbers of induced anagen HFs were significantly reduced both in *K14-Cre*;*Cd44*$^{fl/fl}$ ($n = 6$) (Fig. 4j) and in induced *K14-CreER*^{T};*Cd44*$^{fl/fl}$ ($n = 3$) mice (Fig. 4k). Therefore, the hair growth-activating effect of SPP1 in nevus skin requires epithelial CD44 signalling.

Fig. 4: Effect of SPP1 on hair growth depends on CD44.



a,**b**, Epithelial HF cells in both WT control (**a**) and *Tyr-Nras*^{Q61K} (**b**) mice strongly expressed CD44. Samples were also stained for the epithelial keratin marker KRT14. **c**,**d**, Co-staining for SPP1 and CD44 in *Tyr-Nras*^{Q61K} (**c**) and *Tyr-CreER*^{T2};*Braf*^{V600E} (**d**) skin revealed SPP1^{high} clusters of dermal cells adjacent to CD44^{+} bulge cells with weaker colocalizing SPP1 signal (yellow arrows). **e**, *Cd44*$^{-/-}$ mice showed significantly reduced anagen activation in response to SPP1-soaked beads compared with WT mice.

Representative samples (left) and quantification (right) are displayed. $n = 5$; $P = 0.00938$. **f**, *Cd44*$^{-/-}$ mice showed reduced wound-induced hair growth compared with WT mice. Representative samples (left) and quantification (right) are displayed. $n = 6$ in WT and $n = 5$ in *Cd44*$^{-/-}$; $P = 0.0494$. **g**,**h**, *Tyr-Nras*^{Q61K};*CD44*$^{-/-}$ mice lacking *Cd44* showed rescue of hair cycle quiescence. At P44, *Tyr-Nras*^{Q61K};*Cd44*$^{-/-}$ HFs were in coordinated telogen (**g**). Only rare anagen HFs (arrowheads) were present at P52 (**h**). **i**, Quantification of anagen HFs in *Tyr-Nras*^{Q61K} versus *Tyr-Nras*^{Q61K};*Cd44*$^{-/-}$ mice. Double mutants showed reduced ectopic anagen at P44 and P52. At P44, $n = 12$ and $P = 0.00000000249$; at P56, $n = 12$ and $P = 0.0000166$. **j**,**k**, Both constitutive epithelial-specific *K14-Cre*;*Cd44*$^{fl/fl}$ (**j**) and tamoxifen-induced *K14-CreER*^{T};*Cd44*$^{fl/fl}$ (**k**) mice showed significantly reduced anagen activation in response to SPP1-soaked beads compared with control mice. Representative samples (left) and quantification (right) are displayed. In **j**, $n = 4$ in control and $n = 6$ in mutant; $P = 0.0352$. In **k**, $n = 4$ in control and $n = 3$ in induced mutant; $P = 0.0476$. In **e**,**f**,**i**–**k**, *n* refers to biologically independent samples. *P* values are calculated using unpaired two-tailed Student's *t*-test. $^{*}P \leq 0.05$ and $^{**}P \leq 0.01$. Scale bars, 50 µm (**c**,**d**), 100 µm (**a**,**b**), 200 µm (histology; **g**,**h**), 300 µm (**j**,**k**), 500 µm (**e**,**f**) and 1 mm (wholemount; **g**,**h**).

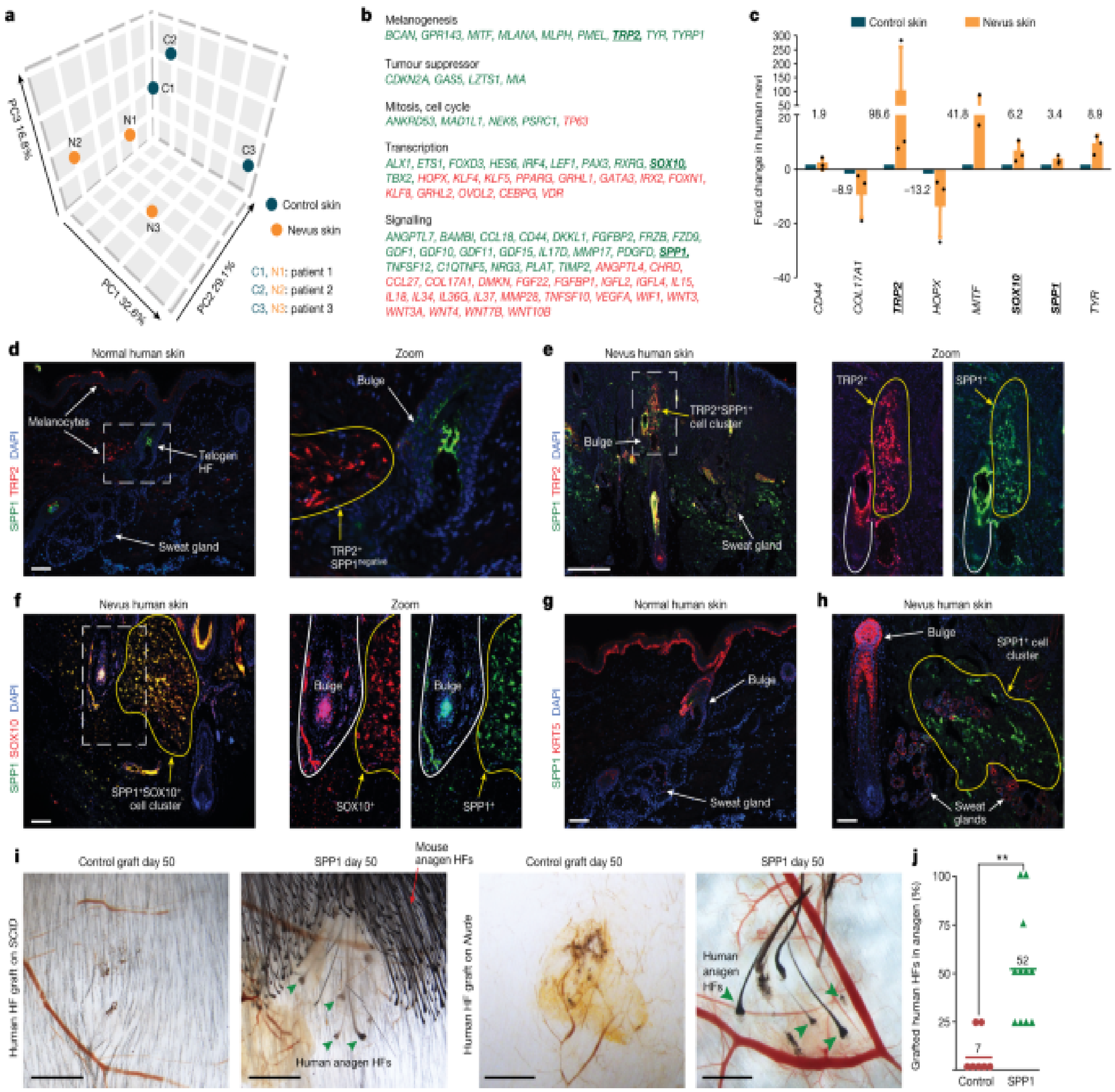
Source data

Human hairy nevi upregulate osteopontin

We also examined signalling aspects of congenital hairy nevi in humans. Whole-tissue RNA-seq revealed prominent differences between congenital hairy nevi and adjacent normal facial skin, and patient-to-patient variability (Fig. 5a, Extended Data Fig. 9n,o and Supplementary Table 5). Nevi showed enrichment for the melanogenesis genes *BCAN*, *GPR143*, *MITF*, *MLANA*, *MLPH*, *PMEL*, *SOX10*, *TRP2*, *TYR* and *TYRP1*, and consistent with *Tyr-Nras*^{Q61K} mouse data, they upregulated the tumour suppressor genes *CDKN2A*, *GAS5*, *LZTS1*, *MIA* and the mitotic markers *ANKRD53*, *MAD1L1*, *NEK6* and *PSRC1*, albeit the latter can be contributed by proliferating HF cells. Among secreted factors, nevi upregulated *SPP1*, several TGFβ/BMP members *GDF1/10/11/15* and *BAMBI*, the WNT modulators *DKKL1*, *FRZB,*

as well as *CCL18*, *IL17D* and *PDGFD* (Fig. 5b). *SPP1* was among upregulated secretome factors shared between human hairy nevi consistently across patients and *Tyr-Nras*^{Q61K} mouse melanocytes (Extended Data Fig. 9p), which we validated by quantitative PCR with reverse transcription (qRT–PCR) (Fig. 5c) and immunostaining (Fig. 5d–h). SPP1 expression was prominent in dermal clusters of either TRP2^{+} (Fig. 5e) or SOX10^{+} melanocytes (Fig. 5f) surrounding bulge regions of HFs. Finally, we tested the hair growth-inducing effect of SPP1 on human scalp HFs in albino *Nude* or pigmented *SCID* host mice[8]. Skin next to telogen HFs were treated with three daily doses of SPP1 or saline. Compared with control ($n = 7$), SPP1 ($n = 11$) accelerated anagen entry in human HFs, sometime accompanied by anagen entry in mouse HFs (Fig. 5i,j). We conclude that SPP1 is a nevus melanocyte-derived hair growth activator in humans.

Fig. 5: Human nevi feature secretome enriched for SPP1.



a, Bulk RNA-seq reveals prominent differences between hairy nevi and adjacent normal facial skin in humans. A principal component analysis plot is shown. See Extended Data Fig. 9. **b**, Selected upregulated (by 2× or more; green) and downregulated (by 2× or more; red) differentially expressed genes in nevus versus normal human skin. Bold and underlined genes were validated by qRT-PCR. **c**, qRT–PCR of selected differentially expressed genes from bulk RNA-seq data. $n = 3$. **d**,**e**, SPP1 and TRP2 co-staining. In normal skin, TRP2^{+} melanocytes did not express SPP1 (**d**), whereas in nevus skin, clusters of TRP2^{+}SPP1^{+} cells were seen next to HF bulge regions (**e**). **f**, SPP1 and SOX10 co-staining. Nevus skin contained SOX10^{+}SPP1^{+} cell clusters next to HF bulge regions. **g**,**h**, SPP1 and KRT5

co-staining. Unlike in normal skin (**g**), SPP1$^+$ cell clusters were seen next to HFs in nevus human skin (**h**). **i**,**j**, SPP1 microinjections induced precocious growth by human scalp HFs (arrowheads). Representative samples of human HFs on day 50 post-grafting (**i**) and quantification of human HFs in anagen (**j**) are shown. In **j**, $n = 7$ for control and $n = 11$ for SPP1; $P = 0.00034$. In **c**, n refers to independent experiments. In **j**, n refers to biologically independent samples. P values were calculated using unpaired two-tailed Student's t-test. $^{**}P \leq 0.01$. Scale bars, 100 μm (**d**–**h**) and 1 mm (**i**).

Source data

Discussion

In this work, we studied how melanocytic skin nevi develop hair overgrowth, which led us to discover that senescent cells can prominently activate tissue-resident SCs and stimulate regeneration. Traditionally, accumulation of senescent cells in tissues is viewed as detrimental to their regenerative potential. This scenario plays out during natural advanced ageing, pathologically accelerated ageing or upon genotoxic exposure[4]. Broad build-up of senescent cells depletes the regenerative capacity of tissues in part via direct elimination of SCs (that is, many SCs become senescent and, thus, non-proliferative) and in part via excessive activation of cytokine-rich secretome (that is, SASP)[17]. SASP factors induce a state akin to low-grade inflammation, which, when persistent, triggers tissue fibrosis. Not surprisingly, systemic depletion of senescent cells in mice delays ageing phenotypes[30], whereas senolytics, drugs that selectively kill senescent cells, have emerged as promising candidate therapeutics for age-related pathologies[31].

However, recent evidence points towards alternative, beneficial effects of senescent cells on tissue growth. Senescent cells form in multiple embryonic tissues, including in the apical ectodermal ridge of the developing limb in mice[18]. Such 'developmental' senescent cells secrete signalling factors thought to instruct growth by surrounding non-senescent embryonic cells. Senescent cells also frequently emerge in non-aged adult tissues upon injury, where SASP factors stimulate enhanced repair. This scenario has been observed in zebrafish after fin amputation[32], in mice following exercise-

induced or cardiotoxin-induced skeletal muscle injury[33,34], surgical resection of liver[35] and excisional skin wounding[20]. In tumours, excessive growth by cancer-initiating cells can rely on stimulating paracrine signals from adjacent senescent cells. The latter can form among cancer-associated stromal cells[21,22] or within cancer cell lineage itself, either triggered by an oncogenic mutation (OIS mechanism) or genotoxic anticancer therapy (DNA damage-induced senescence)[36]. The above examples teach that the paracrine component of the cellular senescence program is commonly used as part of the tissue growth-promoting mechanism. The mechanism of hair overgrowth reported by us in skin nevi exemplifies growth-promoting property of senescent cells (Extended Data Fig. 10). In Supplementary Discussion 1, we discuss conditions necessary for the promoting effect of senescent cells on tissue growth and insights offered by the hairy nevus model.

Whether hairy melanocytic nevus is an outlying example of the kind of effects that senescent cells exert on HFs still remains unknown. Indeed, commonly reduced rather than enhanced hair growth is observed in animal models and in people with increased senescent cell burden—advanced age, progeria or exposure to radiation and chemotherapy. Hair overgrowth is also a leading clinical presentation of smooth muscle hamartoma, a congenital or acquired benign nevus-like condition driven by OIS-activating mutations in cutaneous smooth muscle cells[37]. At the same time, nevus sebaceous, where keratinocytes carry OIS-activating mutations, does not present hair overgrowth, but instead features exuberantly enlarged sebaceous glands[38]. We posit that the exact tissue-level consequence of senescence (for example, hair growth versus sebaceous hypertrophy) depends on the exact molecular composition of SASP, which in turn depends on the original lineage of cells that become senescent, the senescence-inducing mechanism and possibly other factors. That SASP composition is probably heterogenous is also strongly supported by molecular data emerging from other recent studies on the senescent cell secretome (reviewed in Supplementary Discussion 2).

SPP1 is the lead SASP factor secreted by senescent dermal melanocytes that potently induces hair growth. SPP1 is also the topmost SASP factor produced by senescent cancer-associated fibroblasts[22], and its signalling via CD44 promotes cancer cell stemness, tumour growth and radio-resistance[28]. We showed that the hair growth-promoting effect of SPP1 also requires an

intact CD44 receptor on epithelial cells. Consistently, an SPP1 sequence-based synthetic peptide lacking the CD44-binding site fails to promote epithelial proliferation in cultured human HFs[39]. In this context, our data points to future hair growth-stimulating therapies in which select SASP factors, such as SPP1 or its CD44-binding derivatives, are injected into hair loss-affected skin. In support of this approach are clinical cases reporting hair loss-resistant melanocytic nevi on the scalp of patients with alopecia, including alopecia universalis[40].

Several intriguing questions arise from our study that require future investigation. First, not all melanocytic nevi in people are hairy, probably because they do not satisfy all of the conditions necessary for the growth-promoting effect of senescent cells. In-depth comparison of hairy versus non-hairy human nevi will probably reveal new cellular and molecular diversity of these understudied tissue states. Second, in addition to growing more frequently, hairs in human nevi also become thicker and longer, a property known as terminalization. Because hairs in mice cannot undergo terminalization, future studies on human nevus hairs will probably reveal additional signalling effects of SASP on HF cells, beyond SCs. Third, despite carrying activating oncogene mutations, melanocytes in both *Tyr-Nras*^{Q61K} and Tyr-*CreER*^{T2};*Braf*^{V600E} mice become senescent in the dermis next to HFs, but not within HFs themselves. This suggests that a distinct signalling microenvironment within HFs can effectively counteract the OIS mechanism. Future studies comparing signals that melanocytes receive from other cells in their dermal versus HF locations will probably identify new senescence-preventing pathways. Last, normally, melanocytes are not critical regulators of HF SCs and hair growth timing (that is, grey hairs still grow robustly). Thus, acquisition of senescence can confer non-niche cells with novel niche-like properties. By the same accord, acquisition of senescence and SASP by ‘professional’ niche cells (for example, dermal papilla fibroblasts in HFs) may endow them with new regulatory properties. Future works should seek similar effects of cellular senescence on SC functions in other actively renewing organs, such as gut and bone marrow.

In conclusion, our study into the peculiar, yet poorly understood skin condition of hairy nevus led us to identify a distinct regulatory mechanism for adult SCs by tissue-resident senescent cells. These findings have far-

reaching implications for advancing our understanding of SC niche regulation and for developing new therapeutic strategies to regenerative disorders.

Methods

Experimental mouse models

The following mouse lines were used: *Tyr-Nras*^{Q61K}, *Tyr-rtTA*, *Tyr-CreER*^{T2}, *Tyr(C-2J)*, *Braf*^{V600E}, *Trp53*^{flox}, *Spp1*$^{-/-}$, *Spp1*^{flox}, *tetO-Spp1*, *Cd44*$^{-/-}$, *Cd44*^{flox}, *K14-Cre*, *K14-CreER*^{T}, *K14-H2B-GFP*, *K14-Edn3*, *K14-Kitl*, *tdTomato*, *TOPGAL*, *Nude* and *SCID*. Tissue-specific mouse models were produced by crossing either *Cre*-carrying or *CreER*-carrying animals with *flox-ed* gene carrying animals, or *rtTA*-carrying animals with *tetO*-carrying animals. All animal experiments followed all relevant guidelines and regulations and were approved by the Institutional Animal Care and Use Committee at China Agricultural University (to Z.Y.) and/or the Animal Care Committee at Gifu University (to T.K.) and/or the Animal Care and Use Committee of National Taiwan University (to C.-H.K.) and/or the Institutional Animal Care and Use Committee at University of California, Irvine (to B.A. and/or A.K.G. and/or M.V.P.) and/or the Institutional Animal Care and Use Committee at Central South University (to J.L.) and/or the Institutional Animal Care and Use Committee at Kyungpook National University (to J.W.O.).

Mouse induction protocols

Tetracycline-controlled overexpression of SPP1 in melanocytes was achieved in *Tyr-rtTA*;*tetO-Spp1* mice with 2 mg ml^{-1} doxycycline hyclate (Sigma) in 5% sucrose and a doxycycline-containing diet (Bio-Serv, 200 mg kg^{-1}) provided ad libitum. Inducible conditional gene recombination was achieved in *CreER*-carrying and *flox-ed* gene-carrying animals by intraperitoneal injection of tamoxifen (Sigma) in corn oil at a dose of 75 mg kg^{-1}. In P2 animals, inducible conditional gene recombination was achieved by topical administration of (Z)-4-hydroxytamoxifen (4-HT; Sigma) in DMSO at 75 mg ml^{-1}.

EdU pulse and pulse-chase assays

Mice were intraperitoneally injected with EdU (5 μg g^{-1} body weight) daily for seven consecutive days (pulse period), followed by an 8-week chase period. A portion of harvested skin was examined histologically using an EdU imaging kit (Thermo Fisher). Remaining skin portion was used to isolate cells for flow cytometry-based quantification using an EdU flow kit (Thermo Fisher). Triple-positive $CD34^{+}CD49f^{+}EdU^{+}$ cells were used to quantify EdU^{+} bulge SCs.

Protein injection procedure

Intradermal delivery of protein-soaked agarose beads was performed as previously described[8,11]. In brief, recombinant SPP1 protein (441-OP, R&D) was reconstituted in 0.1% BSA to a final concentration of 1.3 mg ml^{-1}. Affi-gel blue beads (Bio-Rad) were washed three times in sterile PBS, air dried and resuspended in reconstituted recombinant protein solution. Beads were incubated on ice for 1 h before implantation. For both recombinant protein and BSA controls, beads were implanted intradermally in P51–P53 animals. Bead implantation sites were resupplied with additional protein at 24, 48 and 72 h.

Skin wounding procedure

Mice were shaved and skin was cleaned with antiseptic. Surgery was conducted under continuous isoflurane anaesthesia. A full-thickness excisional wound was created without injuring the underlying fascia with dermal biopsy punch. Mice were given post-surgical analgesia: subcutaneous ketoprofen, followed by acetaminophen in drinking water.

Flow cytometry and FACS procedures

Dorsal skin was digested into single cells with Dispase II solution (Roche), followed by collagenase I solution (Life Technologies). Cells were filtered first through 70-μM and then 40-μM strainers. Viability dye (BioLegend) was used to exclude dead cells. Cell suspension was stained with primary

antibodies in FACS staining buffer (1% BSA in PBS with 2 mM EDTA) for 30 min on ice before sorting. The following antibodies were used: mouse anti-γH2AX (1:100; 564718, BD Biosciences), mouse anti-TRP2 (1:50; sc-74439 AF647, Santa Cruz Biotechnology), rat anti-Ki67 (1:50; 58-5698-82, Thermo Fisher), rat anti-CD117 (1:100; 105812, BioLegend), rat anti-CD45 (1:50; 103108, BioLegend), rat anti-CD34 (1:50; 560230, BD Biosciences), rat anti-CD49f (1:100; 555736, BD Biosciences) and rabbit anti-SPP1 (1:100; 702184, Thermo Fisher). Cells were sorted on FACSAria II sorters (BD Biosciences) and flow cytometry analysis was performed on LSRII flow cytometer (BD Biosciences). Data were analysed with FlowJo software (version 10.8.0). Expression of SPP1 protein was detected using staining of both permeabilized cells (permeabilized condition) and non-permeabilized cells (surface-bound condition). Under permeabilized condition, we measured total SPP1 present in cells, whereas under surface-bound conditions, we measured SPP1 present on the cell surface, such as bound to its receptors. For permeabilization, cells were washed in PBS and resuspended at 1 million cells per 100 µl, permeabilization buffer was added and cells were stained following Fixation/Permeabilization kit instructions (BD Biosciences).

Primary melanocyte culture assay

Melanocytes were purified from P0 mouse skin by FACS as $CD117^{+}CD45^{neg}$ populations. Sorted cells were then cultured in complete primary melanocyte media (RPMI 1640, 5% FBS, antibiotic–antimycotic, 2.5 ng l^{-1} basic human fibroblast growth factor, 10 µM ethanolamine, 1 mg ml^{-1} insulin, 1 µM *O*-phosphoethanolamine, 5 nM endothelin, 25 nM α-MSH and 50 ng ml^{-1} murine SC factor) at 37 °C with 5% CO_2.

H_2O_2 treatment procedure

Cultured melanocytes in culture dishes or chamber slides were treated with H_2O_2 (Sigma) at 100 mM or vehicle (medium 254 and HMGS-2) for 2 h at 37 °C. Treated cells were rinsed twice with PBS.

DiI labelling procedure

Cells were labelled with DiI dye (Thermo Fisher) following the manufacturer’s instructions. In brief, cells were incubated for 15 min at 37 °C in culture medium supplied with 5 µl of the cell-labelling solution per 1 ml. After labelling, cells were dissociated with Accutase (Stemcell Technologies), followed by two washes with PBS.

Cell injection procedure

Cells were counted using a haemocytometer and then diluted to 2,000 cells per microlitre in cell culture medium. Of cell suspension, 10–50 µl was slowly injected intradermally into the dorsal skin of recipient mice using a 29-G needle.

Grafting procedure

Skin micro-grafts containing four to six anagen HFs were transplanted to the dorsal skin of 6-to-8-week-old female *SCID* or *Nude* mice, as previously described[8]. Thirty days post-grafting, 10 µl of recombinant protein or saline was microinjected to the HF grafting site for 3 consecutive days. Host mice were euthanized on post-grafting day 50 and skin was analysed on wholemount.

ABT-737 treatment procedure

Mice were subcutaneously injected twice (on days P10 and P12) with ABT-737 (Cayman Chemical) or vehicle control at a dose of 75 mg kg^{-1}.

β-Gal staining

For β-galactosidase staining, thick sections (20 µm) were incubated in 1 mg ml^{-1} X-gal substrate in PBS with 1.3 mM $MgCl_2$, 3 mM $K_3Fe(CN)_6$ and 3 mM $K_4Fe(CN)_6$ at 37 °C overnight. For senescence-associated β-gal staining, cells were stained using a kit (Cell Signaling) according to the manufacturer’s instructions. In brief, cells were fixed with fixative solution provided by the manufacturer for 15 min at room temperature, followed by acidic β-gal detection using pH 6.0 staining solution overnight at 37 °C.

Immunohistochemical staining

For paraffin-embedded sections, skin samples were fixed with 4% (vol/vol) paraformaldehyde overnight at 4 °C. Histological sections were permeabilized for 15 min in PBS + 0.1% Triton X-100 (PBST) and blocked for at least 1 h at room temperature with PBST + 3% BSA. Mouse antibodies were blocked with the M.O.M. block kit (Vector Laboratories). Primary antibodies were incubated overnight at 4 °C and secondary antibodies were incubated for 1 h at room temperature. The following primary antibodies were used: rabbit anti-γH2AX (1:300; 9718, Cell Signaling), rabbit anti-TRP2 (1:200; ab74073, Abcam), rabbit anti-TRP2 (1:200; ab103463, Abcam), mouse anti-PCNA (1:1,000; ab29, Abcam), rat anti-CD34 (1:100; 14-0341-82, Thermo Fisher), rabbit anti-SOX9 (1:200; AB5535, Millipore), goat anti-SPP1 (1:100; AF808, R&D), goat anti-SPP1 (1:300; AF1433, R&D), rabbit anti-KRT14 (1:2,000; ab119695, Abcam), rabbit anti-CD44 (1:100; PA5-94934, Thermo Fisher), rabbit anti-SOX10 (1:100; ab180862, Abcam), rabbit anti-KRT5 (1:1,000; 905501, BioLegend) and goat anti-Pcad (1:200; AF761, R&D Systems). The following secondary antibodies were used: donkey anti-rat AF555 (1:1,000; ab150154, Abcam), donkey anti-rabbit AF555 (1:1,000; A31572, Thermo Fisher), donkey anti-mouse AF555 (1:1,000; A31570, Thermo Fisher), donkey anti-rabbit AF488 (1:1,000; A21206, Thermo Fisher), donkey anti-goat AF488 (1:1,000; A11055, Thermo Fisher), goat anti-rat AF488 (1:1,000; A11006, Thermo Fisher), goat anti-rabbit AF488 (1:1,000; 4412s, Cell Signaling), goat anti-mouse AF555 (1:1,000; 4409s, Cell Signaling) and goat anti-rabbit AF555 (1:1,000; 4413s, Cell Signaling).

RNAscope staining

RNA staining was performed using the Multiplex Fluorescent v2 kit (Advanced Cell Diagnostics). In brief, skin was frozen in OCT compound and sectioned at 12–15 µm. Sections were fixed at room temperature for 1 h with 4% paraformaldehyde in PBS, followed by standard manufacturer's protocols (Advanced Cell Diagnostics). RNA probes for hybridization were purchased from Advanced Cell Diagnostics and included *Mm-Spp1* (catalogue no. 435191-C1), *Mm-Dct-C2* (*Trp2*; 460461-C2), *Mm-Cdkn2b*

(*p15*; 458341-C1), *Mm-Cdkn2a* (*p16*; 411011-C1), *Mm-Mki67-C3* (*Ki67*; 416771-C3) and *Mm-Aurkb* (461761-C1).

Western blot assay

Single sorted melanocytes or cells from mouse whole-back skin were lysed in RIPA buffer (Sigma) containing a cocktail of protease inhibitors (Thermo Fisher). Of each cell lysate, 25 µg was loaded onto a 12% separating Bis-Tris gel. Proteins were transferred to a nitrocellulose membrane. Membrane was incubated with primary goat anti-mouse SPP1 antibody (1:100; AF808, R&D) or rabbit anti-β-actin antibody (1:1,000; 4967, Cell Signaling) at a concentration of 2.5 $\mu g\, ml^{-1}$. The blot was developed with Enhanced Chemiluminescence Plus Developer (Fisher Scientific).

ELISA

SPP1 levels in the supernatant of cell cultures were measured by a mouse OPN/SPP1 ELISA kit (Thermo Fisher) according to the manufacturer's instructions. In brief, SPP1 concentration was calculated by generating a standard curve from recombinant SPP1 protein diluted between 0 and 2,000 $pg\, ml^{-1}$. Microplates were measured using a Synergy microplate reader (BIO-TEK) at a wavelength of 450 nm.

Real-time PCR assay

Total RNA from sorted cells was extracted using RNeasy Micro Kit (Qiagen) coupled with its on-column DNase digestion protocol. Total RNA was then reverse-transcribed with Superscript III (Life Technologies) in the presence of oligo-dT. Full-length cDNA was normalized to an equal amount using housekeeping genes *GAPDH* or 18S. Primers are listed in Supplementary Table 6.

Colony-forming assay

Sorted GFP-expressing HF bulge SCs and hair germ progenitors from *K14-H2B-GFP* mice were plated onto 3T3 fibroblast feeder layer cells, pre-treated with mitomycin C to induce cell cycle arrest. Cells were co-cultured

at 37 °C in William's E medium supplemented with calcium and antibiotic–antimycotic. Medium was replaced after 48 h, and the attachment rate was evaluated following an additional 12 h of culture. Attached cells were passaged upon confluence, which was achieved every 4–6 days. Calcium-supplemented culture medium was changed every 2–3 days. In other experiments, bulge SCs were FACS sorted as $CD34^{+}CD49f^{+}$ cells and cultured at a concentration of 1,000 cells per squared centimetre, in the presence of mitomycin C inactivated 3T3 fibroblasts. After 2 weeks, 0.5% crystal violet (Sigma) solution made in 1:1 ratio of water:methanol was added to each culture well. Stained plates were then rinsed with water, air dried and imaged.

Human skin samples

Collection of human skin samples followed all relevant guidelines and regulations and was approved by the Research Ethics Committee at National Taiwan University Hospital and/or the Medical Ethics Committee at Kyungpook National University Hospital and/or the Ethics Committee of Xiangya Hospital, Central South University and comply with guidelines from the Ministry of Science and Technology (MOST) of the People's Republic of China. All participants provided written informed consent. No identifiable images of human research participants are shown.

Bulk and single-cell RNA-seq for mouse tissue

For bulk RNA-seq, total RNA was extracted from FACS-sorted cells in biological triplicates with an RNA integrity number of more than 9.1, and 1 ng of mRNA was used for full-length cDNA synthesis, followed by PCR amplification using Smart-seq2. The libraries were sequenced on the Illumina Next-Seq500 system to an average depth of 10–30 million reads per library using paired 43-bp reads.

For single-cell RNA-seq, cells were captured using the Fluidigm C1 chips as per the manufacturer's protocol. A concentration of 200,000–350,000 cells per millilitre was used for chip loading. After cell capture, chips were examined visually under the microscope to determine the capture rate, and empty chambers or chambers with multiple cells were excluded from the

analysis. cDNA was synthesized and amplified on the Fluidigm C1 Single-Cell Auto Prep System with the Clontech SMARTer Ultra Low RNA kit and the ADVANTAGE-2 PCR kit (Clontech). Single-cell RNA-seq libraries were constructed in 96-well plates according to the Fluidigm C1 manual. Multiplexed libraries were analysed on Agilent 2100 Bioanalyzer for fragment distribution and quantified using Kapa Biosystem's universal library quantification kit. Libraries were sequenced as 75-bp paired-end reads on the Illumina Next-Seq500 platform.

For both bulk and single-cell RNA-seq, reads were first aligned using STAR v.2.4.2a with parameters '--outFilterMismatchNmax 10 --outFilterMismatchNoverReadLmax 0.07 --outFilterMultimapNmax 10' to the reference mouse genome (mm10/genocode,vM8). Gene expression levels were quantified using RSEM v.1.2.25 with expression values normalized into fragments per kilobase of transcript per million mapped reads (FPKM). Samples with more than 1,000,000 uniquely mapped reads and more than 60% uniquely mapping efficiency were used for downstream analyses. Differential expression analysis was performed using edgeR v.3.2.2 on protein-coding genes and long non-coding RNAs. Differentially expressed genes were selected by using fold change ≥ 2, false discovery rate < 0.05 and counts per million reads ≥ 2.

Bulk RNA-seq for human tissue

RNA was extracted from human hairy nevus skin as well as normal skin from nevus edge using the Qiagen RNA extraction kit. cDNA was synthesized using the Superscript III First-strand synthesis system (Invitrogen) and quantified using the Agilent Bioanalyzer. Bulk RNA-seq analysis was performed using the standard pipeline. In brief, pair-end RNA-seq reads were aligned using STAT/2.5.1b to the human reference genome hg38. Gene expression was measured using RESM/1.2/25 with expression values normalized into FPKM.

Single-cell data analysis

For all single-cell data analysis, low-quality cells were filtered out and the same normalization was performed to eliminate cell-specific biases. For

each cell, we calculated three quality control metrics: the number of expressed genes, the total number of transcripts and the proportion of transcripts in mitochondrial genes. The single-cell data matrix was column-normalized (divided by the total number of transcripts and multiplied by 10,000) and then log-transformed with pseudo-count +1.

For single-cell RNA-seq data on bulge SCs, cells from P30 WT, P56 WT and P56 *Tyr-Nras*^{Q61K} samples were combined, and the expression of genes with multiple Ensembl IDs was averaged. For quality control, cells with the total number of TPM counts of less than 750,000, with the proportion of TPM counts in mitochondrial genes of more than 20% and with the number of expressed genes of more than 7,000 or less than 2,000 were removed. In summary, 20 cells were removed, leading to 256 cells for downstream analyses. Clustering of cells was performed using the Seurat R package (V2.3). Principle component analysis (PCA) was first performed using highly variable genes, which were identified with an average expression of more than 0.01 and dispersion of more than 1. We regressed out the effects of the total number of transcripts and the transcripts in mitochondrial genes. The top 17 PCs were selected based on the Jackstraw method (JackStraw function). Using these top PCs, the Louvain modularity-based community detection algorithm was used to obtain cell clusters with resolution being 1.1, giving five clusters. The likelihood-ratio test was used to perform differential gene expression analysis between the clusters. Genes with a *P* value of less than 0.01 and a log fold change greater than 0.25 were considered as differentially expressed. To visualize cells onto a two-dimensional space, we performed *t*-distributed stochastic neighbour embedding. The relatedness of cell clusters was determined by performing unsupervised hierarchical clustering of average gene expression of cell clusters using the highly variable genes (correlation distance metric and average linkage). To determine the cell cycle phase of each cell, we used cell cycle-related genes, including a core set of 43 G1/S and 54 G2/M genes. For each cell, a cell cycle phase (G1, S and G2/M) was assigned based on its expression of these cell cycle-related genes using the CellCycleScoring function in Seurat.

Statistics and reproducibility

Sample size calculations were not performed for mouse experiments, but $n = 3$ is a standard minimal sample size that in our previous studies was found to be sufficient to assess changes in hair growth in mice. Group sizes in animal experiments were derived from the power analysis performed on preliminary experimental data. Animals of both sexes were used, and analyses were not segregated by sex. Age of animals is defined in all experiments in postnatal days. Statistical analyses were performed using unpaired one-tailed or two-tailed (defined in the figure legends) Student's *t*-tests. In all bar charts shown in figures, error bars are mean ± s.d. Statistical significance degree in figures is defined as follows: $P \geq 0.05$ (not significant), $*P \leq 0.05$ and $**P \leq 0.01$; exact *P* values are provided in the figure legends. Differentially expressed gene analysis on RNA-seq data, reported in Supplementary Tables 1, 3, 4 and 5, was done using the edgeR package. When comparing gene expression between groups, the exact test (exactTest() function, two-sided) was performed for *P* value calculation after the negative binomial models were fitted and dispersion was calculated. *P* values were adjusted by using Benjamini and Hochberg's approach for false discovery rate output. For gene ontology terms reported in Supplementary Tables 1, 3, 4 and 5, analysis was done using Metascape. *P* values were calculated using hypergeometric test, and then adjusted by using Benjamini and Hochberg correction. Exact *P* values are reported in the above-mentioned tables. All experiments were repeated independently with similar results of three times or more, and data shown in the figures are from representative experiments. The number of independent repeats for the representative experiments shown as micrographs are as follows: Fig. 1c ($n = 3$), Fig. 1d ($n = 5$), Fig. 1e ($n = 3$), Fig. 1k ($n = 5$), Fig. 2j ($n = 3$), Fig. 3g ($n = 5$), Fig. 4a–d ($n = 3$ each), Fig. 5d–h ($n = 3$ each), Extended Data Fig. 1a,b,d,f–h ($n = 3$ each), Extended Data Fig. 1c ($n = 6$), Extended Data Fig. 1e ($n = 7$), Extended Data Fig. 1j,k ($n = 3$ each), Extended Data Fig. 2a ($n = 4$), Extended Data Fig. 2d ($n = 4$), Extended Data Fig. 3j ($n = 4$), Extended Data Fig. 3k ($n = 4$), Extended Data Fig. 4a ($n = 5$), Extended Data Fig. 4l ($n = 3$), Extended Data Fig. 5a ($n = 6$), Extended Data Fig. 5f–j ($n = 3$ each), Extended Data Fig. 6a,b ($n = 3$ each), Extended Data Fig. 7i ($n = 3$), Extended Data Fig. 8a–c ($n = 4$ each), Extended Data Fig. 8d ($n = 3$), Extended Data Fig. 8e ($n = 5$), Extended Data Fig. 8f ($n = 5$), Extended Data Fig. 8g ($n = 4$), Extended Data Fig. 9e ($n = 3$), Extended Data Fig. 9l,m ($n =$

3 each). Experiments were not randomized or performed in a blinded manner, except where noted.

Schematics

Schematics were prepared using Adobe Illustrator.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Mouse bulk RNA-seq data are located at GSE111999, mouse single-cell RNA-seq data are located at GSE112722 and human bulk RNA-seq data are located at GSE112219. Processed bulk RNA-seq and single-cell data are provided in Supplementary Tables 1–5. Primer sequences are provided in Supplementary Table 6. Source data behind all graphs in main and extended data figures are provided with this paper. Full versions of all gels and blots are provided in Supplementary Fig. 1. Sequential gating strategies are provided in Supplementary Fig. 2. Source data are provided with this paper.

Code availability

No custom code was used. Source data are provided with this paper.

References

1. Scadden, D. T. Nice neighborhood: emerging concepts of the stem cell niche. *Cell* **157**, 41–50 (2014).

2. Pawlikowski, J. S. et al. Wnt signaling potentiates nevogenesis. *Proc. Natl Acad. Sci. USA* **110**, 16009–16014 (2013).

3. Michaloglou, C. et al. BRAF^{E600}-associated senescence-like cell cycle arrest of human naevi. *Nature* **436**, 720–724 (2005).

4. Campisi, J. Aging, cellular senescence, and cancer. *Annu. Rev. Physiol.* **75**, 685–705 (2013).

5. Hsu, Y. C. & Fuchs, E. A family business: stem cell progeny join the niche to regulate homeostasis. *Nat. Rev. Mol. Cell Biol.* **13**, 103–114 (2012).

6. Muller-Rover, S. et al. A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages. *J. Invest. Dermatol.* **117**, 3–15 (2001).

7. Cotsarelis, G., Sun, T. T. & Lavker, R. M. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329–1337 (1990).

8. Liu, Y. et al. Hedgehog signaling reprograms hair follicle niche fibroblasts to a hyper-activated state. *Dev. Cell* https://doi.org/10.1016/j.devcel.2022.06.005 (2022).

9. Wang, Q. et al. A multi-scale model for hair follicles reveals heterogeneous domains driving rapid spatiotemporal hair growth patterning. *eLife* https://doi.org/10.7554/eLife.22772 (2017).

10. Plikus, M. V. et al. Self-organizing and stochastic behaviors during the regeneration of hair stem cells. *Science* **332**, 586–589 (2011).

11. Plikus, M. V. et al. Cyclic dermal BMP signalling regulates stem cell activation during hair regeneration. *Nature* **451**, 340–344 (2008).

12. Festa, E. et al. Adipocyte lineage cells contribute to the skin stem cell niche to drive hair cycling. *Cell* **146**, 761–771 (2011).

13. Ali, N. et al. Regulatory T cells in skin facilitate epithelial stem cell differentiation. *Cell* **169**, 1119–1129.e11 (2017).

14. Chen, C. C. et al. Organ-level quorum sensing directs regeneration in hair stem cell populations. *Cell* **161**, 277–290 (2015).

15. Roh, M. R., Eliades, P., Gupta, S. & Tsao, H. Genetics of melanocytic nevi. *Pigment Cell Melanoma Res.* **28**, 661–672 (2015).

16. Choi, Y. S. et al. Topical therapy for regression and melanoma prevention of congenital giant nevi. *Cell* **185**, 2071–2085.e12 (2022).

17. Coppe, J. P. et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, 2853–2868 (2008).

18. Munoz-Espin, D. et al. Programmed cell senescence during mammalian embryonic development. *Cell* **155**, 1104–1118 (2013).

19. Mosteiro, L. et al. Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science* https://doi.org/10.1126/science.aaf4445 (2016).

20. Demaria, M. et al. An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. Cell* **31**, 722–733 (2014).

21. Wang, T. et al. Senescent carcinoma-associated fibroblasts upregulate IL8 to enhance prometastatic phenotypes. *Mol. Cancer Res.* **15**, 3–14 (2017).

22. Pazolli, E. et al. Senescent stromal-derived osteopontin promotes preneoplastic cell growth. *Cancer Res.* **69**, 1230–1239 (2009).

23. Joost, S. et al. Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. *Cell Syst.* **3**, 221–237.e9 (2016).

24. Dikovskaya, D. et al. Mitotic stress is an integral part of the oncogene-induced senescence program that promotes multinucleation and cell cycle arrest. *Cell Rep.* **12**, 1483–1496 (2015).

25. Rezza, A. et al. Signaling networks among stem cell precursors, transit-amplifying progenitors, and their niche in developing hair follicles. *Cell Rep.* **14**, 3001–3018 (2016).

26. Mori, R., Shaw, T. J. & Martin, P. Molecular mechanisms linking wound inflammation and fibrosis: knockdown of osteopontin leads to rapid repair and reduced scarring. *J. Exp. Med.* **205**, 43–51 (2008).

27. Joost, S. et al. Single-cell transcriptomics of traced epidermal and hair follicle stem cells reveals rapid adaptations during wound healing. *Cell Rep.* **25**, 585–597.e7 (2018).

28. Pietras, A. et al. Osteopontin–CD44 signaling in the glioma perivascular niche enhances cancer stem cell phenotypes and promotes aggressive tumor growth. *Cell Stem Cell* **14**, 357–369 (2014).

29. Miletti-Gonzalez, K. E. et al. Identification of function for CD44 intracytoplasmic domain (CD44-ICD): modulation of matrix metalloproteinase 9 (MMP-9) transcription via novel promoter response element. *J. Biol. Chem.* **287**, 18995–19007 (2012).

30. Baker, D. J. et al. Clearance of p16^{Ink4a}-positive senescent cells delays ageing-associated disorders. *Nature* **479**, 232–236 (2011).

31. Robbins, P. D. et al. Senolytic drugs: reducing senescent cell viability to extend health span. *Annu. Rev. Pharmacol. Toxicol.* **61**, 779–803 (2021).

32. Da Silva-Alvarez, S. et al. Cell senescence contributes to tissue regeneration in zebrafish. *Aging Cell* **19**, e13052 (2020).

33. Young, L. V. et al. Muscle injury induces a transient senescence-like state that is required for myofiber growth during muscle regeneration. *FASEB J.* **36**, e22587 (2022).

34. Saito, Y., Chikenji, T. S., Matsumura, T., Nakano, M. & Fujimiya, M. Exercise enhances skeletal muscle regeneration by promoting senescence in fibro-adipogenic progenitors. *Nat. Commun.* **11**, 889 (2020).

35. Cheng, N., Kim, K. H. & Lau, L. F. Senescent hepatic stellate cells promote liver regeneration through IL-6 and ligands of CXCR2. *JCI Insight* https://doi.org/10.1172/jci.insight.158207 (2022).

36. Lee, S. & Schmitt, C. A. The dynamic nature of senescence in cancer. *Nat. Cell Biol.* **21**, 94–101 (2019).

37. Adulkar, S. A., Dongre, A. M., Thatte, S. S. & Khopkar, U. S. Acquired smooth muscle hamartoma. *Indian J. Dermatol. Venereol. Leprol.* **80**, 483 (2014).

38. Groesser, L. et al. Postzygotic HRAS and KRAS mutations cause nevus sebaceous and Schimmelpenning syndrome. *Nat. Genet.* **44**, 783–787 (2012).

39. Alam, M. et al. An osteopontin-derived peptide inhibits human hair growth at least in part by decreasing FGF7 production in outer root sheath keratinocytes. *Br. J. Dermatol.* https://doi.org/10.1111/bjd.18479 (2019).

40. Yamamoto, T., Okabe, H. & Hoshi, M. Alopecia totalis sparing congenital melanocytic nevus: Renbök phenomenon. *Dermatol. Sin.* **37**, 176–177 (2019).

Acknowledgements

M.V.P. is supported by LEO Foundation grants LF-AW-RAM-19-400008 and LF-OC-20-000611, Chan Zuckerberg Initiative grant AN-0000000062, W. M. Keck Foundation grant WMKF-5634988, NSF grant DMS1951144 and NIH grants U01-AR073159, R01-AR079470, R01-AR079150, R21-AR078939 and P30-AR075047. Additional support comes from the NSF grant DMS1763272 and Simons Foundation grant (594598 to Q.N.), California Institute for Regenerative Medicine SRL grant CL1-00520-1.2, an unrestricted gift from S. Segura. B.A. is supported by R01-AR044882. A.K.G. is supported by U54-CA217378 and R01-CA244571. J.D.E. is supported by R61/33-AR073031. Xiaojie Wang is supported by a CIHR post-doctoral fellowship (MFE-123724) and UC Irvine NSF-Simon Center for Multiscale Cell Fate Research fellowship. E.N.G. is supported by the

NSF GRFP grant DGE-1321846. J.L. is supported by the Training Program of the Major Research Plan of the National Natural Science Foundation of China (91749114), the National Natural Science Foundation of China (81773351) and the National Natural Science Funds for Distinguished Young Scholars (82225039). H.-L.L. is supported by the NIH training grant T32-CA009054. Y.L. is supported by the Chinese Government scholarship and by the clinical fellowship from the California Institute for Regenerative Medicine training grant (EDUC4-12822). The image in Fig. 1a is reproduced with permission from S. Liber.

Author information

Authors and Affiliations

1. Department of Developmental and Cell Biology, University of California, Irvine, CA, USA

   Xiaojie Wang, Raul Ramos, Kosuke Yamaga, Shan Jiang, Chen-Hsiang Kuan, Truman Kt Nguyen, Heidi Y. Liang, Nitish Udupi Shettigar, Renzhi Hou, Kevin H. Tran, Andrew Nguyen, Kimberly N. Vu, Jennie L. Phung, Jonard P. Ingal, Katelyn M. Levitt, Xiaoling Cao, Yingzi Liu, Christian F. Guerrero-Juarez, Rabi Murad, Hye-Lim Lee, Ali Mortazavi, Qing Nie & Maksim V. Plikus

2. Sue and Bill Gross Stem Cell Research Center, University of California, Irvine, CA, USA

   Xiaojie Wang, Raul Ramos, Kosuke Yamaga, Chen-Hsiang Kuan, Truman Kt Nguyen, Heidi Y. Liang, Nitish Udupi Shettigar, Renzhi Hou, Kevin H. Tran, Andrew Nguyen, Kimberly N. Vu, Jennie L. Phung, Jonard P. Ingal, Katelyn M. Levitt, Xiaoling Cao, Yingzi Liu, Vanessa M. Scarfone, Guangfang Wang, Kara Nicole Paolilli, Xiaoyang Wang, Christian F. Guerrero-Juarez, Lily Halida Putri Widyastuti, Hye-Lim Lee, Bogi Andersen & Maksim V. Plikus

3. NSF-Simons Center for Multiscale Cell Fate Research, University of California, Irvine, CA, USA

Xiaojie Wang, Raul Ramos, Renzhi Hou, Christian F. Guerrero-Juarez, Bogi Andersen, Qing Nie & Maksim V. Plikus

4. Glycobiology Research and Training Center, Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA, USA

Anne Q. Phan & Jeffrey D. Esko

5. Department of Biological Chemistry, University of California, Irvine, CA, USA

Jessica L. Flesher, Sohail Jahid, Elyse Noelani Greenberg, Rolando Ruiz-Vega, Priya Vasudeva, Bogi Andersen & Anand K. Ganesan

6. Center for Complex Biological Systems, University of California, Irvine, CA, USA

Shan Jiang, Christian F. Guerrero-Juarez, Rabi Murad, Ali Mortazavi, Qing Nie & Maksim V. Plikus

7. Department of Anatomy, Yonsei University College of Medicine, Seoul, Korea

Ji Won Oh

8. Department of Anatomy and Hair Transplantation Center, Kyungpook National University and Hospital, Daegu, Korea

Ji Won Oh

9. Department of Mathematics, University of California, Irvine, CA, USA

Suoqin Jin, Christian F. Guerrero-Juarez & Qing Nie

10. School of Mathematics and Statistics, Wuhan University, Wuhan, China

Suoqin Jin

11. Division of Plastic Surgery, Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan

Chen-Hsiang Kuan

12. Research Center for Developmental Biology and Regenerative Medicine, National Taiwan University, Taipei, Taiwan

Chen-Hsiang Kuan

13. Amplifica Holdings Group, Inc., San Diego, CA, USA

Nitish Udupi Shettigar & Maksim V. Plikus

14. Department of Dermatology, Xiangya Hospital, Central South University, Changsha, China

Yingzi Liu, Zhili Deng & Ji Li

15. Department of Tissue and Organ Development, Regeneration and Advanced Medical Science, Gifu University Graduate School of Medicine, Gifu, Japan

Nobuhiko Taguchi & Takahiro Kunisada

16. Department of Physiology and Biophysics, University of California, Irvine, CA, USA

Ryan T. Davis & Devon A. Lawson

17. Centre for Skin Sciences, University of Bradford, Bradford, UK

Kevin J. McElwee

18. University of Bordeaux, INSERM U1034, Adaptation cardiovasculaire à l'ischémie, Pessac, France

Alain-Pierre Gadeau

19. Department of Medicine, University of California, Irvine, CA, USA

Bogi Andersen

20. State Key Laboratory of Farm Animal Biotech Breeding, College of Biological Sciences, China Agricultural University, Beijing, China

Zhengquan Yu

21. Laboratory of Gene Regulation and Signal Transduction, Departments of Pharmacology and Pathology, University of California San Diego, School of Medicine, La Jolla, CA, USA

Michael Karin

22. Institute for Comparative Molecular Endocrinology (CME), University of Ulm, Helmholtzstrasse 8/1, Ulm, Germany

Jan Tuckermann

23. Leibniz Institute on Aging-Fritz Lipmann Institute, Beutenbergstrasse 11, Jena, Germany

Jan Tuckermann

24. Department of Dermatology, University of California, Irvine, CA, USA

Anand K. Ganesan

25. Hunan Key Laboratory of Aging Biology, Xiangya Hospital, Central South University, Changsha, China

Ji Li

Contributions

M.V.P. is the lead corresponding author on this paper and oversaw the project. Xiaojie Wang and M.V.P. wrote the manuscript and generated the figures. Xiaojie Wang, R.R., A.Q.P., T.K.N., N.U.S., H.Y.L., A.N., K.N.V., J.L.P., K.H.T., K.M.L., N.T., K.N.P., C.F.G.-J., L.H.P.W., X.C., J.W.O. and H.-L.L. carried out transgenic animal experiments. Xiaojie Wang, S. Jiang, S. Jin, R.T.D., R.M. and J.L. carried out RNA-seq analyses. Xiaojie Wang, Xiaoyang Wang, J.L.F., V.M.S. and G.W. carried out the cell culture and colony-forming assays. Xiaojie Wang, K.Y., E.N.G., R.R.-V., S. Jahid and P.V. carried out western blot and protein assays. Xiaojie Wang, J.P.I., C.-H.K. and R.H. performed the RNAscope assays. Y.L., Z.D., J.W.O., C.-H.K. and J.L. carried out the human assays. A.-P.G., T.K., M.K. and J.T. provided critical advice. M.V.P., J.W.O., C.-H.K., D.A.L., Z.Y., T.K., J.D.E., A.K.G. and J.L. supervised experiments. M.V.P. and Q.N. supervised the RNA-seq data analysis. K.J.M., D.A.L., B.A., A.M., Q.N., T.K., J.D.E. and A.K.G. edited the paper.

Corresponding authors

Correspondence to Xiaojie Wang or Maksim V. Plikus.

Ethics declarations

Competing interests

M.V.P. is an inventor on a patent application filed by the University of California, Irvine describing the use of senescent secretome factors for promoting hair growth, among other claims. M.V.P. is also a co-founder and Chief Scientific Officer at the Amplifica Holdings Group, Inc. N.U.S. is employed by the Amplifica Holdings Group, Inc. All other authors declare no competing interests.

Peer review

Peer review information

Nature thanks Maria Kasper and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Nevus mouse models exhibit ectopic hair growth.

a–h, At all postnatal time points examined, *Tyr-Nras*^{Q61K} mice showed ectopic anagen HFs. In WT control mice, HFs are in first anagen at day P15 (**a**); in first telogen at P23 (**b**); in second anagen at P36 (**c**); in extended second telogen at P44 (**d**), P56 (**e**), P62 (**f**), P69 (**g**); and in third telogen at P100 (**h**). In contrast, at all of the above time points, *Tyr-Nras*^{Q61K} skin contained many ectopic anagen HFs (green arrowheads). For each time point, representative wholemount (left) and histology samples (right) are shown. **i**, Schematic representation of the hair cycle state in WT control mice (top) and *Tyr-Nras*^{Q61K} mice (bottom) at indicated time points (middle). Colors: green – anagen, yellow – catagen, red – telogen. **j**, Albino *Tyr-Nras*^{Q61K} mice (crossed onto an albino *Tyr(C-2J)* background) maintain ectopic hair growth phenotype (green arrowheads) both at P56 (left) and P100 (right). For each time point, representative wholemount and histology samples are shown. **k**, *Tyr-CreER*^{T2}*;Braf*^{V600E} mice induced with tamoxifen at P2-4 formed nevi and exhibited ectopic hair growth. At all postnatal time points examined (P44, P69 and P100), dorsal skin in induced *Tyr-CreER*^{T2}*;Braf*^{V600E} mice contained many ectopic anagen HFs (green arrowheads). In contrast, HFs in induced control mice at the above time points were in telogen. For each time point, representative wholemount and histology samples (for *Tyr-CreER*^{T2}*;Braf*^{V600E} mice) are shown. Also, see

Extended Data Fig. 4a,b. Scale bars, **a–h, j** (wholemount) – 1 mm; **k** (wholemount) – 500; **a–h, j, k** (histology) – 200 μm.

Extended Data Fig. 2 Nevus melanocytes but not normal melanocytes are sufficient to activate hair growth.

a, Compared to skin from control mice, skin from P69 *Tyr-CreER^{T2};Braf^{V600E}* mice induced with tamoxifen at P21-25, contained clusters of *Trp2^{+}/p15^{+}/Ki67^{neg}* melanocytes in upper dermis next to bulge regions of HFs. **b, c**, Compared to skin from control mice, skin from P69 *Tyr-CreER^{T2};Braf^{V600E}* mice induced at P21-25, contained significantly more TRP2^{+}/Ki67^{neg} melanocytes (**b**) and TRP2^{+}/γH2AX^{+} melanocytes (**c**). Left – representative cytometry plots, right – quantification. In **b**, $n = 3$; $P = 0.00239$. In **c**, $n = 3$; $P = 0.00675$. **d**, Compared to skin from control mice (left), skin from *Tyr-Nras^{Q61K}* mice (center) and *Tyr-CreER^{T2};Braf^{V600E}* mice (right) contained clusters of *Trp2^{+}/p16^{+}/Ki67^{neg}* melanocytes in upper dermis next to bulge regions of HFs. In **b, c**, *n* = independent experiments. *P* values are calculated using unpaired two-tailed Student's t-test. $^{**}P \leq 0.01$. Scale bars, **d** – 50 μm; **a** – 100 μm.

Source data

Extended Data Fig. 3 Nevus melanocytes and senescent melanocytes stimulate new hair growth.

a–i, Melanocyte grafting experiments, in which melanocyte lineage cells were isolated as tdTomato^{+} cells from mice that contain *Tyr-CreER^{T2}* and *tdTomato* constructs, and that were induced with tamoxifen at P21. **a–e**, tdTomato^{+} melanocytes were isolated from P56 *Tyr-Nras^{Q61K};Tyr-CreER^{T2};tdTomato* (**a**) and P69 *Tyr-CreER^{T2};Braf^{V600E};tdTomato* nevus mouse skin (**c**) and injected into *SCID* mouse skin. Both types of mutant melanocytes (**b, e**) induced anagen after 21 days. **f–i**, Control melanocyte lineage cells were isolated from *Tyr-CreER^{T2};tdTomato* mice during telogen at P56 (**f**) and during anagen at P33 (**h**) and intradermally injected into telogen skin of *SCID* mice. Cells from both conditions (**g, i**) did not induce ectopic anagen 21 days after injection. Representative samples are shown in

(**b, e, g, i**). Anagen HFs for experiments from (**b, e, g, i**) are quantified in (**d**). In **d**, *n* = 4. **j–o**, H_2O_2-induced senescence experiment (**j**). Senescent status of H_2O_2-treated melanocytes was confirmed with senescent β-Gal staining (**k, l**). H_2O_2-treated (**n**), but not control DiI-labeled melanocytes (**m**) induced anagen 21 days after injection into *SCID* mice. Anagen HFs are quantified in (**o**). In **l**, *n* = 4; *P* = 0.0112. In **o**, *n* = 7 in vehicle, *n* = 6 in H_2O_2; *P* = 0.000024. In **d, l, o**, *n* = biologically independent samples. *P* values are calculated using unpaired two-tailed Student's t-test. $^*P \leq 0.05$, $^{**}P \leq 0.01$. Scale bars, **m, n** – 1 mm; **a, b, c, e, f–i** – 2 mm; **j, k** – 200 µm.

Source data

Extended Data Fig. 4 Induction of nevi results in hair growth hyper-activation and osteopontin overexpression.

a-b, Compared to control, *Tyr-CreER^{T2};Braf^{V600E}* mice induced with tamoxifen at P2-4, showed prominent hair growth. Representative P56 skin samples are shown in (**a**). Anagen HFs are marked in **(a)** and quantified in (**b**). In **b**, at P44 (*n* = 12 in control, *n* = 20 in mutant; *P* = 0.00218), at P56 (*n* = 12 in control, *n* = 21 in mutant; *P* = 0.00000000804), at P69 (*n* = 12 in control, *n* = 16 in mutant; *P* = 0.0000526), at P100 (*n* = 12 in control, *n* = 16 in mutant; *P* = 0.00000662). **c, d**, Following tamoxifen-induction at P21-25, *Tyr-CreER^{T2};Braf^{V600E}* mice developed nevi by P44 and started to display ectopic hair growth from P56 onward. Representative wholemount and histology samples at five time points between P44-100 are shown in (**c**), and anagen HF density quantification is shown in (**d**). In **d**, *n* = 9. **e**, On ELISA, SPP1 levels became significantly higher in the supernatant from day 5 cultures of primary sorted *Tyr-CreER^{T2};Braf^{V600E}* melanocytes at five indicated timepoints from P56 onward. Data from P69 cells is also shown in main Fig. 3e. In **e**, at P44 (*n* = 4; *P* = 0.2686), at P56 (*n* = 4; *P* = 0.0000269), at P62 (*n* = 4; *P* = 0.003095), at P69 (*n* = 4; *P* = 0.00224), at P100 (*n* = 4; *P* = 0.0035). **f-i**, On cytometry, SPP1 levels in permeabilized cells (**f, h**) as well as surface-bound SPP1 levels in non-permeabilized cells (**g, i**) were significantly higher in melanocytes from induced *Tyr-CreER^{T2};Braf^{V600E}* mice (pBraf, sBraf) compared to wild type control mice (pWT, sWT) at indicated time points. Representative cytometry plots are shown in (**f, g**)

and quantification is show in (**h, i**). Data from P69 cells is also shown in main Fig. 3d. In **h**, at P44 ($n = 3$ for pWT, $n = 4$ for pBraf; $P = 0.000318$), at P56 ($n = 3$; $P = 0.0000533$), at P62 ($n = 3$; $P = 0.00426$), at P69 ($n = 3$; $P = 0.001397$), at P100 ($n = 3$ for pWT, $n = 4$ for pBraf; $P = 0.00000386$). In **i**, at P44 ($n = 3$; $P = 0.0531$), at P56 ($n = 3$; $P = 0.0912$), at P62 ($n = 3$; $P = 0.2495$), at P69 ($n = 3$; $P = 0.291$), at P100 ($n = 3$; $P = 0.00399$). **j–k**, Cytometry of permeabilized (**j**) and non-permeabilized melanocytes (**k**) showed significantly higher levels of SPP1 compared to isotype control both in *Tyr-Nras*^{Q61K} mice (pNras, sNras) and *Tyr-CreER*^{T2}*;Braf*^{V600E} mice (pBraf, sBraf). Representative cytometry plots are shown on the left and quantification on the right of **j** and **k**. In **j**, for pNras ($n = 3$; $P = 0.00000175$), for pBraf ($n = 3$; $P = 0.00000213$). In **k**, for sNras ($n = 3$; $P = 0.00297$), for sBraf ($n = 3$; $P = 0.000000536$). **l**, Skin of P69 *Tyr-CreER*^{T2}*;Braf*^{V600E} mice contained *Trp2*$^{+}$*/Spp1*$^{+}$ melanocytes in upper dermis adjacent to bulge regions of HFs. In **b, d**, n = biologically independent samples. In **e, h, i, j, k**, n = independent experiments. Data are mean ± s.d. P values are calculated using unpaired one-tailed (in **b** at P56) or two-tailed (in **b** at P44, P69, P100, **e, h, i, j, k**) Student's t-test. NS, $P \geq 0.05$, *$P \leq 0.05$, **$P \leq 0.01$. Scale bars, **a** (wholemount) – 1 mm; **c** (wholemount) – 300 μm; **c** (histology) – 200 μm; **a** (histology), **l** – 100 μm.

Source data

Extended Data Fig. 5 Effect of ABT-737 treatment and non-nevus melanocyte expansion on hair cycle.

a-b, Unlike vehicle, subcutaneous ABT-737 treatment of *Tyr-Nras*^{Q61K} mice at P10 and P12 decreased fur pigmentation and reduced anagen HFs at P56 (**a**). Anagen HFs are quantified in (**b**). In **b**, n=21; $P = 0.0000454$. **c-e**, Effect of ABT-737 treatment on melanocytes, bulge stem cells and hair cycle status. **c**, On cytometry at P56, the percentage of TRP2^{+}/Annexin V^{+} melanocytes in *Tyr-Nras*^{Q61K} mice significantly increased in response to ABT-737 treatment at P10-12. In **c**, $n = 5$; $P = 0.0001816$. **d**, On cytometry at P56, the abundance of CD34^{+}/CD49f^{+} bulge stem cells in *Tyr-Nras*^{Q61K} mice was unchanged by ABT-737 treatment at P10-12. In **d**, $n = 5$; $P = 0.7891838$. **e**, ABT-737 treatment at P10-12 did not affect normal anagen

timing in WT mice – skin from both vehicle and ABT-737 treated animals contained HFs in anagen at P33. In **e**, $n = 7$; $P = 0.2898739$. In (**c, d, e**) representative data is shown on the left, and data quantification – on the right. **f–j**, Mice with non-nevus expansion in melanocytes display normal hair cycle timing. **f–i**, Similar to control mice, *K14-Edn3* mice with dermal melanocyte expansion (**f, g**) and *K14-Kitl* mice with epidermal melanocyte expansion (**h, i**) were in synchronized anagen at P36 (**f, h**) and synchronized telogen at P56 (**g, i**). **j**, After tamoxifen induction at P12-14, *Tyr-CreER;p53$^{fl/fl}$* mice with melanocyte-specific deletion of *p53*, did not form nevi and exhibited telogen HFs at P56, analogous to induced control mice. In **b**, *n* = biologically independent samples. In **c, d, e** *n* = independent experiments. Data are mean ± s.d. *P* values are calculated using unpaired two-tailed Student's t-test. NS, $P \geq 0.05$, $^{*}P \leq 0.05$, $^{**}P \leq 0.01$. Scale bars, **f–j** (wholemount) – 1 mm; **a, f–j** (histology) – 200 μm; **e** – 100 μm.

Source data

Extended Data Fig. 6 Gene expression patterns in *Tyr-Nras^{Q61K}* HFs stem cells.

a–b, *Tyr-Nras^{Q61K}* telogen HFs maintain normal expression patterns of bulge and hair germ markers. Co-immunostaining for CD34 (red) and Pcad (green) showed that their expression pattern in *Tyr-Nras^{Q61K}* telogen HFs (**b**) is consistent with that in WT telogen HFs (**a**). Bulge cells are CD34^{+} and Pcad^{low}, while hair germ (HG) cells are CD34^{neg} and Pcad^{high}. **c–d**, RNA-seq analysis on WT and *Tyr-Nras^{Q61K}* bulge SCs. Venn diagrams and DEGs heatmap are shown in (**c**). Venn diagrams identify 21 downregulated and 104 upregulated genes specific to *Tyr-Nras^{Q61K}*. Bubble charts in (**d**) show enriched (red) and depleted (green) GO terms in P56 WT bulge SCs, and enriched GO terms in *Tyr-Nras^{Q61K}* bulge SCs (blue). **e–j**, Gene expression patterns in bulge stem cells on single-cell RNA-seq. Cell clusters are color-coded according to main Fig. 2d and are as follows: C1 – inner bulge cells, present in P30 and P56 WT samples and in P56 *Tyr-Nras^{Q61K}* sample; C2 – anagen-specific outer bulge cells, present in P30 WT and P56 *Tyr-Nras^{Q61K}* samples; C3 – telogen-specific outer bulge cells, present in P56 WT sample; C4 and C5 – outer bulge cells specific to *Tyr-*

Nras^{Q61} sample. Violin plots are shown with normalized expression values along the Y-axis. **e**, Expression patterns of outer bulge markers, showing enrichment in clusters C2, C3 and C4. **f**, Expression patterns of telogen-phase outer bulge markers, showing enrichment in cluster C3. **g**, Expression patterns of inner bulge markers, showing enrichment in cluster C1. **h**, Expression patterns of anagen-phase enriched markers. **i**, Expression patterns of mutant-enriched markers in cluster C4. **j**, Expression patterns of mutant-enriched markers in cluster C5. Scale bars, **a, b** – 100 µm.

Extended Data Fig. 7 Characterization of bulge stem cells and melanocytes in nevus mouse models.

a, b, Labeling efficiency of bulge stem cells after 7 days of EdU pulse was consistent between WT control and *Tyr-Nras*^{Q61K} mice. Representative cytometry plots (left) and data quantification (right) are shown for all CD34^{+}/CD49f^{+} bulge stem cells in (**a**) and for their CD34^{+}/CD49f^{+}/EdU^{+} subset in (**b**). In **a**, $n = 3$; $P = 0.63929124$. In **b**, $n = 3$; $P = 0.23549636$. **c-f**, On RNA-seq, P56 *Tyr-Nras*^{Q61K} melanocytes differ from P30 and P56 WT melanocytes. PCA is shown in (**c**) and DEGs heatmap in **(d)**. Secretome factors upregulated in *Tyr-Nras*^{Q61K} melanocytes and *Tyr-Nras*^{Q61K}/WT fold change values are shown in (**e**). Bubble chart showing enriched GO terms in *Tyr-Nras*^{Q61K} melanocytes are shown in (**f**). Selected bubbles are colored and annotated. **g**, qRT-PCR validation of selected differentially expressed genes from bulk RNA-seq data on melanocytes. In **g**, $n = 3$. **h**, Venn diagram showing the degree of overlap between the P56 *Tyr-Nras*^{Q61K} melanocyte transcriptome and published *in vitro* human senescent melanocyte secretome (blue, 68%) and core *in vitro* SASP factors (green, 71%). **i**, Compared to skin from wild type control mice (left), skin from *Tyr-Nras*^{Q61K} mice (middle) and induced *Tyr-CreER*^{T2}*;Braf*^{V600E} mice (right) contained clusters of *Trp2*$^{+}$*/Aurkb*$^{+}$ melanocytes in upper dermis next to bulge regions of HFs. In **a, b**, n = independent experiments. Data are mean ± s.d. P values are calculated using unpaired two-tailed Student's t-test. NS, $P \geq 0.05$. Scale bars, **i** – 100 µm.

Source data

Extended Data Fig. 8 Changes in expression and the effect of osteopontin deletion in *Tyr-Nras*^{Q61K} skin.

a-c, Osteopontin expression is increased in *Tyr-Nras*^{Q61K} skin. **a**, *Spp1* reporter activity was increased in *Tyr-Nras*^{Q61K} skin. LacZ staining (blue) on *Tyr-Nras*^{Q61K}*;Spp1*$^{+/-}$ *vs*. control *Spp1*$^{+/-}$ P56 reporter mouse skin showed broad increase in LacZ^{+} cells. Dermal and dermal papilla (DP) expression sites are marked. For each panel, representative wholemount and histology samples are shown on the left and on the right, respectively. **b–c**, Co-immunostaining for KRT5 (red) and SPP1 (green) in P56 WT control (**b**) and *Tyr-Nras*^{Q61K} skin **(c)**. *Tyr-Nras*^{Q61K} skin showed prominently increased SPP1 expression in the dermal compartment, including around bulge regions of HFs (inserts). **d-g**, Osteopontin deletion rescues hair cycle quiescence in *Tyr-Nras*^{Q61K} mice. *Tyr-Nras*^{Q61K}*;Spp1*$^{-/-}$ mice showed rescue of hair cycle quiescence. Unlike *Tyr-Nras*^{Q61K} mice (see Extended Data Fig. 1), *Tyr-Nras*^{Q61K}*;Spp1*$^{-/-}$ mice showed synchronized catagen at P18 (**d**), synchronized anagen at P30 (**e**), synchronized telogen at P44 (**f**) and largely synchronized telogen P52 (**g**). For each time point, representative *Spp1*$^{-/-}$ control and *Tyr-Nras*^{Q61K}*;Spp1*$^{-/-}$ mutant skin samples are shown. Wholemount samples are shown on the right and histology on the left of each panel. Scale bars, **b, c** – 100 μm; **a, d–g** (wholemount) – 500 μm; **a, d–g** (histology) – 200 μm.

Extended Data Fig. 9 Expression and the effect of *Cd44* deletion on hair growth.

a-e, Expression of *Cd44* and related genes. **a**, Relative abundance of *Cd44* isoforms established from full-length bulk RNA-seq. Isoforms are numbered using conventional nomenclature and indicated along the X-axis. *Cd44v* isoforms are designated with “v” and *Cd44s* isoform with “s”. Skin cell types are listed along the Y-axis. WT – wild type cells, MUT – *Tyr-Nras*^{Q61K} mutant cells. Bulge stem cells are enriched for *Cd44v* isoforms 201, 202, 205 and 208. **b**, *Cd44* is prominently expressed on bulk RNA-seq in bulge stem cells from WT control (blue) and *Tyr-Nras*^{Q61K} mice (orange) both at P30 and P56. *Mmp9*, direct downstream target of CD44-ICD

signalling is prominently overexpressed in *Tyr-Nras*^{Q61K} bulge stem cells both at P30 and P56. **c**, Expression values of γ-secretase complex genes in P56 WT and *Tyr-Nras*^{Q61K} bulge stem cells. **d**, Expression values of transcription factors mediating CD44-ICD signalling in P56 WT and *Tyr-Nras*^{Q61K} bulge stem cells. Average FPKM values are shown on **b–d**. In **b-d**, $n = 2$. **e**, LacZ staining (blue) in *Tyr-Nras*^{Q61K}*;Cd44*$^{+/-}$ *vs*. control *Cd44*$^{+/-}$ P56 reporter mice showed LacZ^{+} cells in the skin, both in epithelial and dermal compartments. For each panel, wholemount and histology samples are shown on the left and on the right, respectively. **f-m**, Effects of *Cd44* deletion on bulge stem cells and hair cycle status in nevus mice. **f**, Total abundance of CD34^{+}/CD49f^{+} bulge stem cells remained unchanged in germline *Cd44*$^{-/-}$ mice *vs*. WT control mice. In **f**, $n = 3$; $P = 0.52$. **g**, Labeling efficiency of bulge stem cells after 7 days of EdU pulse remained unchanged in germline *Cd44*$^{-/-}$ mice *vs*. WT control mice. In **g**, $n = 3$; $P = 0.401$. In (**f, g**) representative cytometry plots are shown on the left, and data quantification on the right. **h**, In *in vitro* culture assay on FACS-isolated bulge stem cells, clonogenic potential of CD34^{+}/CD49f^{+} cells remained unchanged in germline *Cd44*$^{-/-}$ mice *vs*. WT control mice. Top – representative culture plates, bottom – data quantification. In **h**, $n = 6$; $P = 0.384$. **i**, Total abundance of CD34^{+}/CD49f^{+} bulge stem cells remained unchanged in epithelial-specific conditional *K14-Cre;Cd44*$^{fl/fl}$ (aka *Cd44*$^{fl/fl}$) mice *vs*. control mice. In **i**, $n = 3$; $P = 0.328$. **j**, Labeling efficiency of bulge stem cells after 7 days of EdU pulse remained unchanged in *Cd44*$^{fl/fl}$ mice *vs*. control mice. In **j**, $n = 3$; $P = 0.218$. In (**i, j**) representative cytometry plots are shown on the left, and data quantification on the right. **k**, In *in vitro* culture assay on FACS-isolated bulge stem cells, clonogenic potential of CD34^{+}/CD49f^{+} cells remained unchanged in *Cd44*$^{fl/fl}$ mice *vs*. control mice. Top – representative culture plates, bottom – data quantification. In **k**, $n = 6$; $P = 0.411$. **l, m**, *Tyr-Nras*^{Q61K}*;Cd44*$^{-/-}$ mice showed rescue of hair cycle quiescence. Unlike *Tyr-Nras*^{Q61K} mice (see Extended Data Fig. 1), *Tyr-Nras*^{Q61K}*;Cd44*$^{-/-}$ animals showed synchronized anagen at P30 (**l**) and only very occasional ectopic anagen HFs at P56 (**m**). For both time points, representative *Cd44*$^{-/-}$ control and *Tyr-Nras*^{Q61K}*;Cd44*$^{-/-}$ mutant skin samples are shown. Wholemount samples are shown on the right and histology on the left of each panel. **n–o**,

Bulk RNA-seq between hairy nevi and adjacent normal facial skin in humans. DEGs heatmap is shown in (**n**), bubble chart with enriched (blue) and depleted (orange) GO terms in human nevi is shown in (**o**). **p**, Comparisons of human nevus secretome (blue) with *Tyr-Nras*^{Q61K} mouse melanocyte secretome (red) and published in vitro SASP (yellow). In **b–d, f–k**, *n* = independent experiments. Data are mean ± s.d. *P* values are calculated using unpaired two-tailed Student's t-test. NS, $P \geq 0.05$. Scale bars, **e, l, m** (wholemount) – 500 µm; **e, l, m** (histology) – 200 µm.

Source data

Extended Data Fig. 10 Model of senescent cell-induced hair growth in skin nevus.

a, Limited accumulation of senescent cells adjacent to normal, intact stem cell niche can augment it and result in stem cell activation. Mechanism, effects, and examples are listed below the schematic drawing. Tissue stem cells – orange (quiescent) and green (activated); normal niche cells – blue; senescent cells – purple. **b**, Schematic representation of the mechanism driving hair growth hyperactivation in skin nevus. Dermal clusters of senescent melanocytes (purple) secrete SASP factors (colored geometric shapes). SPP1 (blue squares) is the leading SASP factor of senescent melanocytes. It signals via CD44 receptor (black Y shape) to epithelial stem cells in adjacent hair follicles, inducing them into precocious growth (green arrow).

Supplementary information

Supplementary Information

This file contains Supplementary Discussion, Supplementary Figures 1 &2 and Supplementary References

Reporting Summary

Supplementary Table 1

Differential gene expression between sorted bulge stem cells in *Tyr-Nras^{Q61K}* mutant *vs*. control mice as revealed by bulk RNA-sequencing.

Supplementary Table 2

Single-cell RNA-sequencing gene expression matrix for sorted bulge stem cells from *Tyr-Nras^{Q61K}* mutant mice at postnatal day P56 and control mice at postnatal days P30 and P56.

Supplementary Table 3

Differential gene expression between bulge stem cell clusters as revealed by combined single-cell RNA-sequencing analysis of *Tyr-Nras^{Q61K}* mutant and control mouse cells.

Supplementary Table 4

Differential gene expression between sorted dermal melanocytes in *Tyr-Nras^{Q61K}* mutant *vs*. control mice as revealed by bulk RNA-sequencing.

Supplementary Table 5

Differential gene expression between nevus *vs*. non-nevus control human skin as revealed by bulk RNA-sequencing.

Supplementary Table 6

Sequence of mouse-specific and human-specific primers used for RT-PCR.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Fig. 5

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 9

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Wang, X., Ramos, R., Phan, A.Q. *et al.* Signalling by senescent melanocytes hyperactivates hair growth. *Nature* **618**, 808–817 (2023). https://doi.org/10.1038/s41586-023-06172-8

- Received: 11 September 2019
- Accepted: 05 May 2023
- Published: 21 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06172-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06172-8

- Article
- Open Access
- Published: 31 May 2023

In situ tumour arrays reveal early environmental control of cancer immunity

- Guadalupe Ortiz-Muñoz,
- Markus Brown,
- Catherine B. Carbone,
- Ximo Pechuan-Jorge,
- Vincent Rouilly,
- Henrik Lindberg,
- Alex T. Ritter,
- Gautham Raghupathi,
- Qianbo Sun,
- Tess Nicotra,
- Shreya R. Mantri,
- Angela Yang,
- Jonas Doerr,
- Deepti Nagarkar,
- Spyros Darmanis,
- Benjamin Haley,
- Sanjeev Mariathasan,
- Yulei Wang,
- Carlos Gomez-Roca,
- Carlos Eduardo de Andrea,
- David Spigel,
- Thomas Wu,
- Lelia Delamarre,
- Johannes Schöneberg,

- Zora Modrusan,
- Richard Price,
- Shannon J. Turley,
- Ira Mellman &
- …
- Christine Moussion

Show authors

Nature **volume 618**, pages 827–833 (2023)

- 12k Accesses
- 26 Altmetric
- Metrics details

Abstract

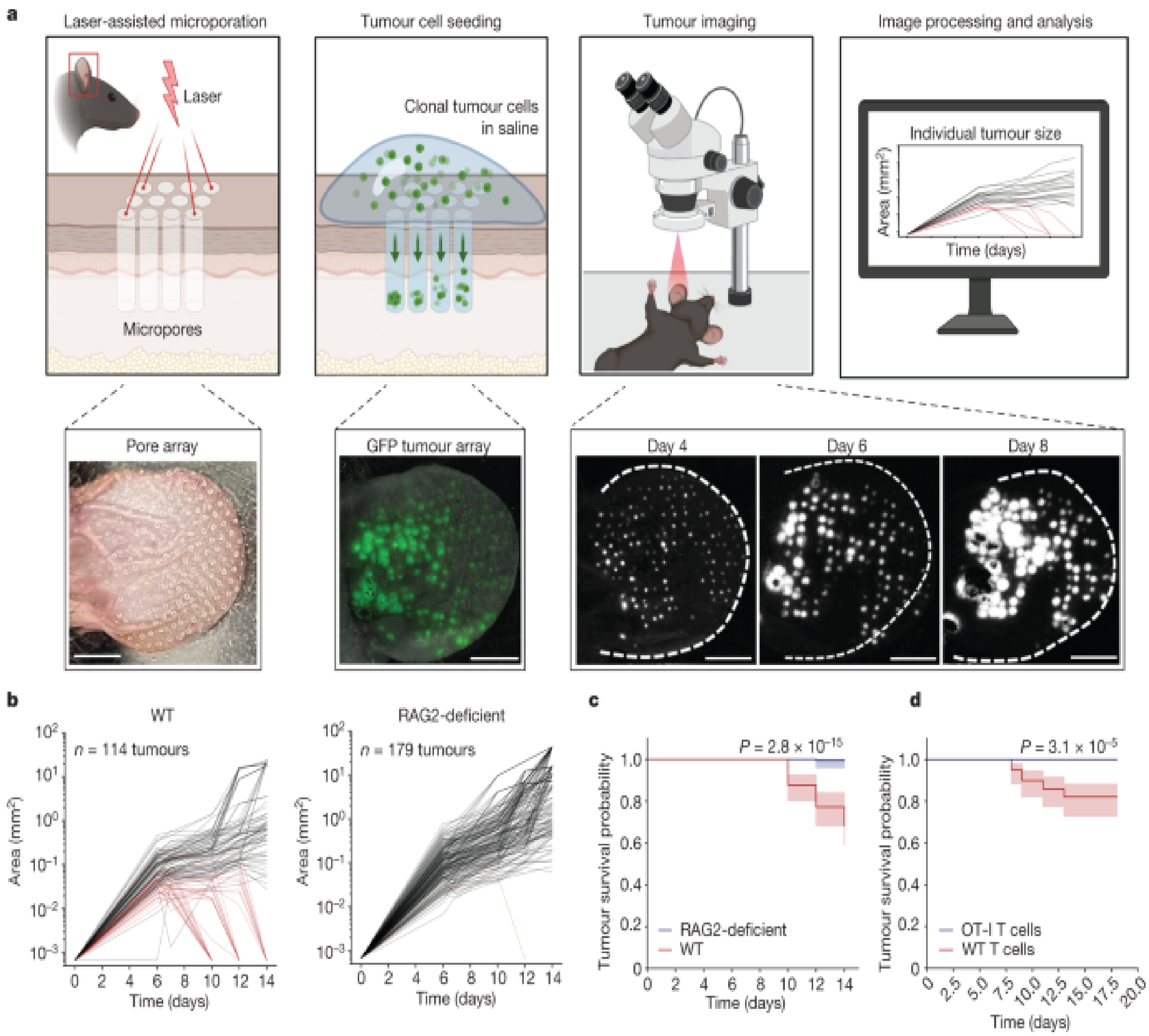
The immune phenotype of a tumour is a key predictor of its response to immunotherapy[1,2,3,4]. Patients who respond to checkpoint blockade generally present with immune-inflamed[5,6,7] tumours that are highly infiltrated by T cells. However, not all inflamed tumours respond to therapy, and even lower response rates occur among tumours that lack T cells (immune desert) or that spatially exclude T cells to the periphery of the tumour lesion (immune excluded)[8]. Despite the importance of these tumour immune phenotypes in patients, little is known about their development, heterogeneity or dynamics owing to the technical difficulty of tracking these features in situ. Here we introduce skin tumour array by microporation (STAMP)—a preclinical approach that combines high-throughput time-lapse imaging with next-generation sequencing of tumour arrays. Using STAMP, we followed the development of thousands of arrayed tumours in vivo to show that tumour immune phenotypes and outcomes vary between adjacent tumours and are controlled by local factors within the tumour microenvironment. Particularly, the recruitment of T cells by fibroblasts and monocytes into the tumour core was supportive of T cell cytotoxic activity and tumour rejection. Tumour immune phenotypes were dynamic over time

and an early conversion to an immune-inflamed phenotype was predictive of spontaneous or therapy-induced tumour rejection. Thus, STAMP captures the dynamic relationships of the spatial, cellular and molecular components of tumour rejection and has the potential to translate therapeutic concepts into successful clinical strategies.

Main

The STAMP technique uses an infrared laser[9] to create an array of hundreds of pores in the dermis of the mouse ear. Tumour cells expressing a fluorescent reporter are seeded into each pore and tumour growth is monitored over the next 4–6 weeks using live fluorescence microscopy (Fig. 1a and Extended Data Fig. 1a–d). To track and measure the growth of dozens of individual microtumours in a single array, we developed a computational pipeline using a convolutional neural network (Extended Data Fig. 1e–h). STAMP can successfully implant orthotopic (melanoma) or heterotopic (mammary, pancreas, lung and colon carcinoma) tumour cell lines in the skin.

Fig. 1: STAMP reveals local T-cell-mediated rejection of clonal skin tumour array.



a, The STAMP workflow. Skin microporation using the P.L.E.A.S.E. laser device and subsequent seeding of tumour arrays from cell suspension. Individual tumours of the array are longitudinally tracked using epifluorescence microscopy and the growth kinetics are analysed by automated computation. Scale bars, 4 mm. **b**, Automated analysis of the growth kinetics of individual KPP-eGFP tumours (area in mm^2) inside the same tumour array implanted into RAG2-deficient (*n* = 179 tumours, *n* = 4 mice) or wild-type (WT) (*n* = 114 tumours, *n* = 3 mice) mice. Sampling is representative of *n* = 5 mice per group. The red lines indicate tumours that were rejected, and the grey lines indicate tumours that persist. **c**, The survival probability of individual tumours of KPP-eGFP arrays as described in **b**. The centre line shows the Kaplan–Meier curve and the shaded area shows the 95% confidence interval. Statistical analysis was performed using

a log-rank test. **d**, The survival probability of individual tumours of KPP-eGFP arrays implanted into RAG2-deficient mice reconstituted by adoptive transfer of tdTomato^{+} T cells from either WT mice ($n = 100$ tumours, $n = 2$ mice) or OT-I mice ($n = 88$ tumours, $n = 2$ mice). Sampling is representative of $n = 2$ independent experiments, $n = 7$ mice per group. The centre line shows the Kaplan–Meier curve and the shaded area shows the 95% confidence interval. Statistical analysis was performed using log-rank tests.

Source Data

Local T cell-mediated tumour rejection

To characterize immune infiltration into STAMP tumours, we compared a clonal KPP-eGFP cell line (pancreatic ductal adenocarcinoma)[10] injected subcutaneously into the flank or using STAMP on the ear. We observed similar immune infiltrates at 10 days after tumour implantation, regardless of the implantation method (Extended Data Fig. 1i). Importantly, immune cell profiling using flow cytometry demonstrated a specific tumour-dependent recruitment of lymphoid and myeloid populations (Extended Data Fig. 1j), and the inflammation associated with laser microporation completely resolved a few days after tumour implantation (Extended Data Fig. 1k–l).

To assess the role of adaptive immunity in controlling tumour growth, KPP-eGFP tumour arrays were implanted into wild-type or immunodeficient RAG2-deficient mice. Wild-type and RAG2-deficient mice displayed comparable initial tumour burdens (Extended Data Fig. 2a); however, after 14 days, immunocompetent mice demonstrated local rejection of around 30% of individual tumours from the same array, while immunodeficient mice did not reject the tumours (Fig. 1b,c and Extended Data Fig. 2b). Tumour survival analysis in wild-type and CD8-depleted mice confirmed that the anti-tumour immune response in STAMP involves CD8 T cells (Extended Data Fig. 2c). To assess the role of antigen-specific CD8 T cells in this local rejection, we implanted KPP-eGFP tumours in RAG2-deficient mice and reconstituted the mice with naive tdTomato^{+} T cells either from mice that contain a polyclonal population of T cells, or OT-I mice, which contain a monoclonal population of ovalbumin-specific CD8 T cells (Supplementary Video 1). In contrast to mice reconstituted with polyclonal

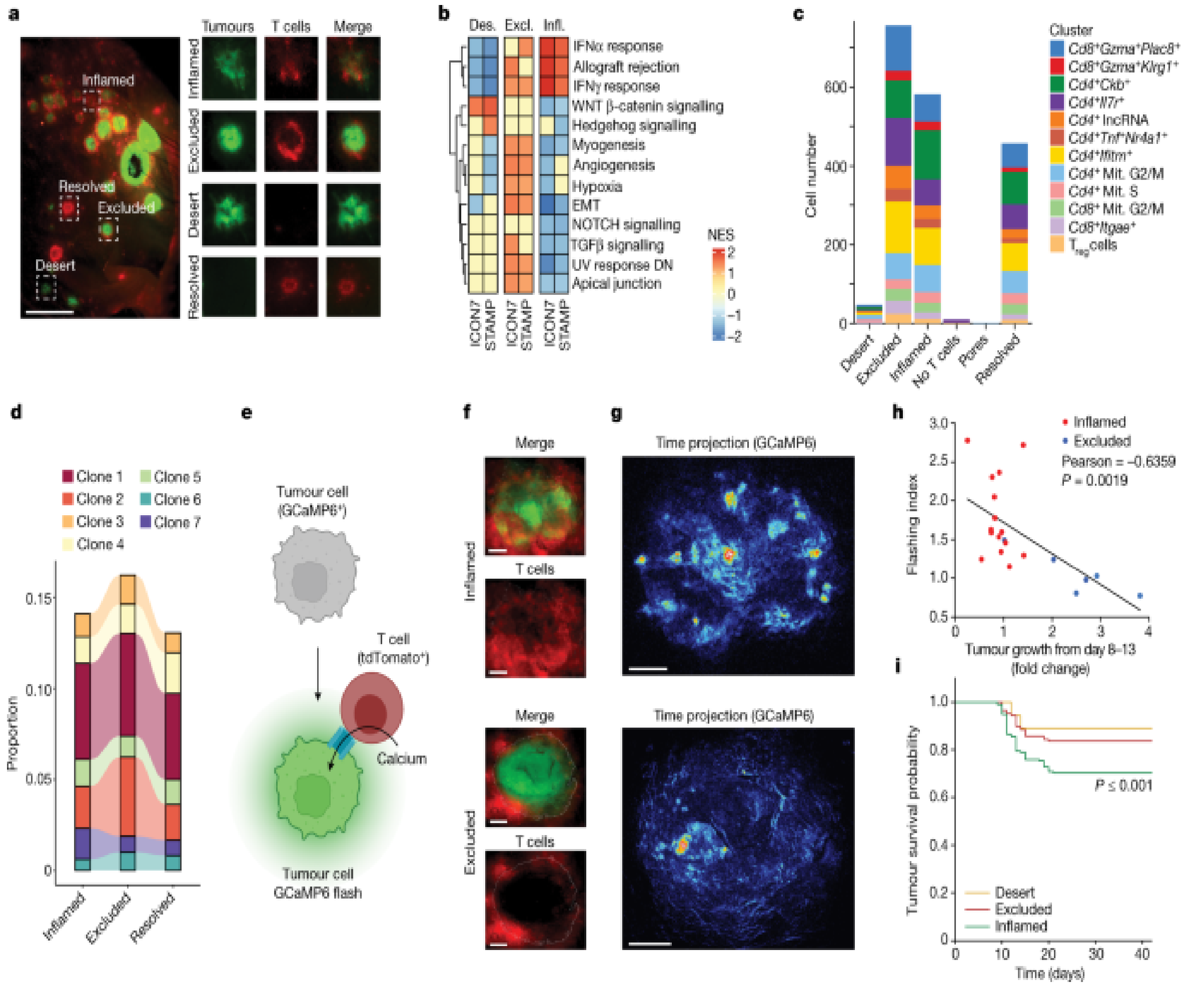
T cells, antigen-mismatched T cells from OT-I mice did not mediate tumour rejection and were not substantially recruited to tumour sites (Fig. 1d and Extended Data Fig. 2d–f). As the KPP-eGFP tumour cells in this experiment do not express ovalbumin, we can conclude that bystander T cells were not sufficient to promote rejection of STAMP tumours.

These observations were corroborated by the converse experiment in which adoptive transfer of antigen-specific T cells was examined in mice that were implanted with KPP tumours expressing a model antigen (M86 tumour antigen or ovalbumin). As summarized in Extended Data Fig. 2g–j, reconstitution with antigen-specific CD8 T cells resulted in heightened T cell recruitment and local tumour rejection. Thus, our findings support a role for antigen-specific T cells recruited into STAMP tumours in mediating local rejection.

Clinical relevance of STAMP phenotypes

To further elucidate the drivers of local tumour rejection, we characterized the spatial distribution of T cells in individual tumours of the STAMP array using fluorescence microscopy. Despite being derived from a clonal tumour cell line, STAMP tumour array exhibited a combination of immune-inflamed, immune-excluded and immune-desert tumours at all of the analysed timepoints (Fig. 2a and Extended Data Fig. 3c). Moreover, we noted a fourth late-onset phenotype, termed resolved tumour, in which $eGFP^{+}$ tumour cells disappeared leaving behind a cluster of $tdTomato^{+}$ T cells.

Fig. 2: Immune-inflamed phenotype supports T cell effector function and tumour rejection.



a, Representative image of a STAMP array of KPP-eGFP tumours at 8 days after tumour implantation in RAG2-deficient mice, reconstituted with tdTomato^{+} T cells. $n = 50$ mice, $n = 10$ independent experiments. Red, T cells; green, KPP-eGFP cells. Left, representation of the entire ear. Right, magnified images of individual tumours with different immune phenotypes. Scale bar, 2 mm. **b**, Heat map comparing the normalized enrichment scores for pathways that are significantly enriched across the immune-inflamed (infl.), immune-desert (des.) and immune-excluded (excl.) phenotypes from either human tumours from the ICON7 clinical trial or mouse STAMP tumours. Normalized enrichment scores were determined using clusterProfiler::GSEA using the false-discovery rate *P*-value adjustment method; $P_{adj} < 0.2$ was considered to be significant. **c**, The abundance of T cell subsets was determined using scRNA-seq analysis of STAMP tumour biopsies pooled by immune phenotype. T_{reg} cells, regulatory T cells. Mit., mitotic. **d**, The relative abundance of seven dominant T cell clonotypes

across immune phenotypes. **e**, Schematic of cytotoxic T cell attack creating calcium-permeable pores in the tumour cell membrane, which triggers green fluorescence of the GCaMP6 calcium sensor in the tumour cell. **f**, Representative images of KPP-mTagBFP2-GCaMP6 STAMP tumours with the inflamed (top) or excluded (bottom) immune phenotype. $n = 6$ tumours. Red, T cells; green, GCaMP6. **g**, Time projection GCaMP6 fluorescent flashes of the tumour described in **f**. **h**, Correlation analysis of the GCaMP6 flashing index at 8 days after tumour implantation and the tumour growth fold change between day 8 and day 13 for immune-inflamed and immune-excluded tumours described in **f**. Pearson correlation was computed assuming a normal distribution. Statistical analysis was performed using two-tailed *t*-tests. **i**, Kaplan–Meier curve showing the survival probability of individual tumours of KPP-eGFP arrays that were immune phenotyped as immune-desert, immune-excluded or immune-inflamed by imaging 8 days after tumour implantation. $n = 632$ tumours, $n = 10$ mice. Statistical analysis was performed using a log-rank test (referenced to excluded tumours). For **f** and **g**, scale bars, 100 μm (inflamed) and 200 μm (excluded).

Source Data

The ability of STAMP to establish different tumour immune phenotypes (TIPs) provided the opportunity to characterize gene expression and pathways associated with each tumour niche. STAMP arrays were first imaged to classify tumours by immune phenotype and individual tumours were then isolated by punch biopsy for bulk RNA-sequencing (RNA-seq) analysis. Immune-desert, immune-excluded and immune-inflamed tumours were found to be transcriptionally distinct (Fig. 2b, Extended Data Fig. 3a and Supplementary Table 1). For each TIP, the patterns of gene expression in mouse STAMP tumours were notably similar to those observed in human tumours (ovarian, ICON7 phase III; bladder, IMvigor210 phase II) (Extended Data Fig. 3b, Supplementary Data 1 and Supplementary Table 2). Consistent with recent reports[11,12,13,14,15], inflamed tumours were enriched for IFNα, IFNγ and allograft rejection signatures, and were also characterized by a downregulation of the TGFβ and NOTCH pathways[11,14]. Immune-excluded tumours were enriched for epithelial–mesenchymal transition (EMT), angiogenesis, hypoxia and ultraviolet light response signatures, suggesting enrichment in the stromal and myofibroblast

components[12,13]. Finally, immune-desert tumours were enriched for WNT/β-catenin and Hedgehog pathways and showed a downregulation of immune signatures[14,15]. The similarity of gene signatures associated with the three TIPs across different human cancer types (ovarian, bladder and lung) and mouse STAMP tumours suggests that STAMP may recapitulate the general mechanisms underlying TIPs observed clinically.

To test whether the coexistence of different immune phenotypes is conserved across multiple tumour models and tissues in mice, we compared different tumour cell lines implanted with STAMP on the ear and on the flank, and in an experimental lung metastasis model (Extended Data Fig. 3c–f and Supplementary Video 2). We observed that, in each tumour model, neighbouring tumours could exhibit disparate phenotypes, including immune-inflamed, immune-excluded and immune-desert tumours. Notably, the relative frequencies of each phenotype varied depending on the cancer cell line used, confirming the importance of tumour cells themselves in influencing the probability of developing a certain TIP.

Although the tumours in STAMP arrays shared the same clonal origin, it was possible that the different TIPs reflected rapid genetic divergence after implantation. To address this issue, we performed whole-exome sequencing (Extended Data Fig. 4a–c) and single-cell RNA-sequencing (scRNA-seq; Extended Data Fig. 4d–f) analysis of tumour cells before implantation and 3 weeks after implantation. As shown in Extended Data Fig. 4, neither pre-existing heterogeneity in the parental cell line nor genetic drift of the individual tumour cells growing in vivo correlated with the TIP diversity observed in STAMP. Thus, our findings indicate that clonally derived tumour cells have the unanticipated ability to develop heterogeneous immune phenotypes, demonstrating that the appearance of a given immune phenotype is not completely dependent on either host or tumour genetics.

Inflamed niches support T cell function

We next examined whether the development of distinct immune phenotypes in a single tumour array reflected qualitative differences in T cell responses between neighbouring lesions. Although tumour-specific T cells in a single ear were probably derived from the same draining lymph nodes, differences

in T cell infiltration and tumour growth control might have reflected differences in T cell receptor clonotype profiles. To this end, we performed 5′ scRNA-seq and T cell receptor sequencing analysis of individual tumours collected from the same mouse and compared T cells from immune desert, excluded, inflamed and rejected tumours (Fig. 2c,d). We identified 12 clusters of CD4 and CD8 T cells, including several clusters of effector T cells, T resident memory cells, regulatory T cells and mitotic cells (Fig. 2c, Extended Data Fig. 5a,b and Supplementary Table 3).

Notably, we did not observe any difference in the absolute or relative abundance of T cell subsets between TIPs on the basis of scRNA-seq or flow cytometry analysis (Fig. 2c and Extended Data Fig. 5c,d). Furthermore, we identified seven immunodominant clonotypes, the relative and absolute abundance of each was similar across TIPs (Fig. 2d and Extended Data Fig. 5e,f).

Although there were no differences in the overall abundance of T cell subsets or clonotypes, the transcriptional profiles of individual T cell clones between immune phenotypes exhibited differentially enriched pathways (Extended Data Fig. 5g, Supplementary Data 2 and Supplementary Table 4). Perhaps reflecting their participation in an ongoing immune response, T cell clones present in inflamed tumours were characterized by increased transcriptional and translational activity as well as mitochondrial biogenesis compared to the same T cells clones located in excluded tumours.

As mitochondrial dynamics influence T cell fate[16,17,18] and translational activity correlates with the activation and differentiation states of effector T cells[19], we investigated the functionality of the T cells present in immune-inflamed versus immune-excluded tumours. We engineered a clonal KPP tumour cell line expressing GCaMP6 calcium sensor[20] that emits green fluorescence after effector T cell attack (Fig. 2e). We validated our experimental approach by demonstrating that calcium flashing occurs as a result of a T cell attack in organoids in vitro (Extended Data Fig. 6a–e and Supplementary Videos 3 and 4) and in vivo (Extended Data Fig. 6f,g and Supplementary Videos 5–7). In a RAG2-deficient model with and without T cell reconstitution, we show that T cells are required for calcium flashing in tumour cells (Extended Data Fig. 6h,i).

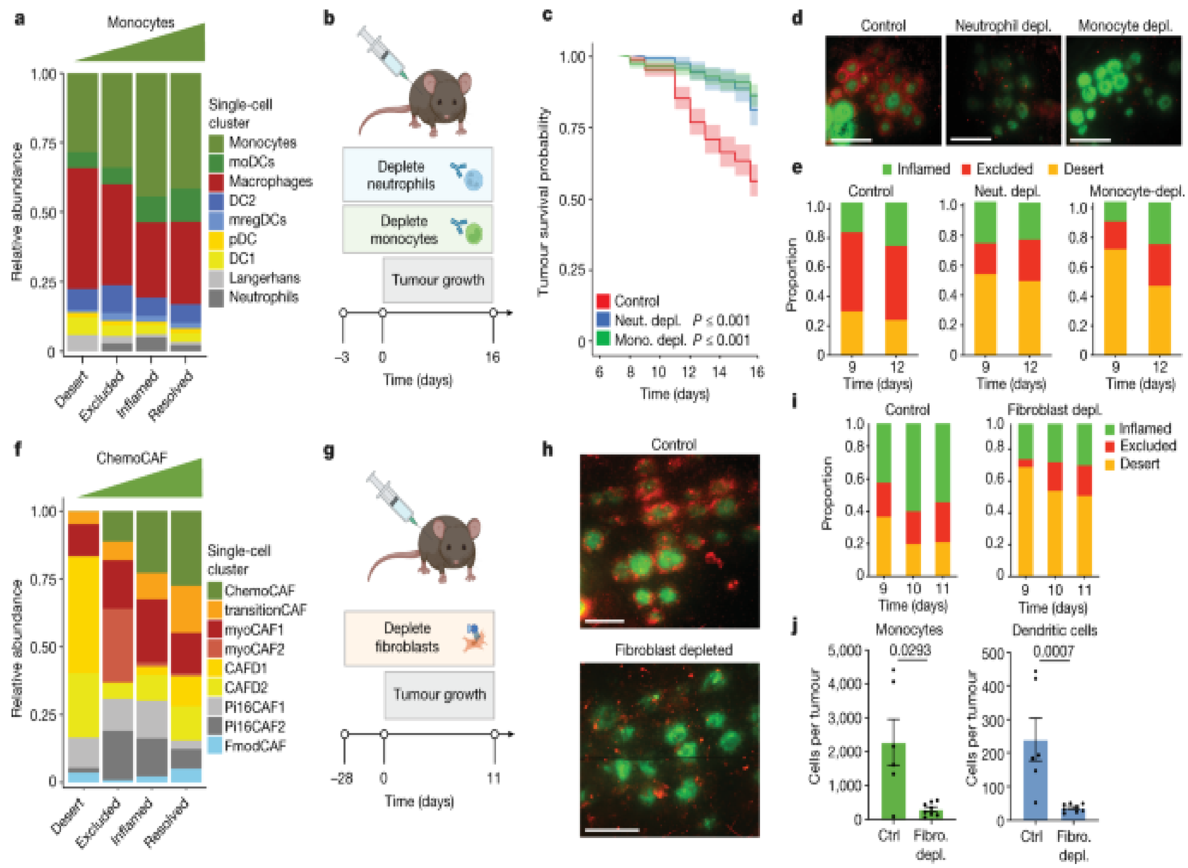
We next quantified differences in calcium flashing between STAMP tumours within a single array. Higher flashing indices were found in inflamed tumours (Fig. 2f–h, Extended Data Fig. 6j–m and Supplementary Videos 8 and 9) and corresponded to a slow tumour growth rate (Fig. 2h). Thus, inflamed tumours that nurture effector T cell function were more likely to regress compared with immune-excluded or immune-desert tumours (Fig. 2i).

Taken together, these results indicate that T cells of the same TCR clonotype exhibit an improved functional capacity when localized in immune-inflamed tumours, emphasizing a determinative role for the tumour microenvironment in shaping the activity and fate of endogenous effector T cells.

Myeloid–stroma control of immune phenotype

To examine the role of the tumour microenvironment in the development of TIPs, we created a scRNA-seq atlas of STAMP microtumours at early timepoints after T cell infiltration (Extended Data Fig. 7a), focusing on myeloid (Extended Data Fig. 7b and Supplementary Table 5) and stromal cell subsets (Extended Data Fig. 8a and Supplementary Table 6). We found that monocytes and monocyte-derived cells were primary producers of the T cell chemoattractants CXCL9 and CXCL10 (Extended Data Fig. 7c) and that increased monocyte and neutrophil abundances were associated with the inflamed and resolved TIPs (Fig. 3a); similar results were obtained both by flow cytometry and deconvolution of bulk RNA-seq (Extended Data Fig. 7d,e). We demonstrated a functional role for these myeloid cell subsets in the recruitment and spatial patterning of T cells by depleting Ly6C^{+} monocytes or Ly6G^{+} neutrophils (Fig. 3b and Supplementary Data 3). Notably, depletion of either myeloid population resulted in a significant decrease in tumour rejection (Fig. 3c), which was accompanied by a strong decrease in overall T cell recruitment and an increase in desert tumours (Fig. 3d,e and Extended Data Fig. 7f,g).

Fig. 3: Myeloid and stromal cells control TIP and tumour fate.



a, The relative abundance of myeloid cell subclusters as determined by scRNA-seq analysis of STAMP tumour biopsies pooled by immune phenotype. DCs, dendritic cells; moDCs, monocyte-derived dendritic cells; mregDCs, mature DCs with immunoregulatory molecules; pDCs, plasmacytoid dendritic cells. **b**, The experimental design relating to **c**–**e**. Neutrophils or monocytes were ablated using depleting antibodies (Gr1 or Ly6C, respectively) beginning 3 days before STAMP implantation of KPP-eGFP tumour arrays in E8I CD8-cre LSL-tdTomato immunocompetent mice. $n = 6$ isotype-control-treated mice, $n = 319$ tumours; $n = 5$ neutrophil-depleted (neut. depl.) mice, $n = 187$ tumours; $n = 5$ monocyte-depleted (mono. depl.) mice, $n = 301$ tumours. **c**, The survival probability of individual tumours of KPP-eGFP arrays related to **b**. The centre line shows the Kaplan–Meier curve and the shaded area shows the 95% confidence interval. Statistical analysis was performed using log-rank tests (referenced to the isotype control). **d**, Representative image of STAMP tumour arrays of non-depleted (left) neutrophil-depleted (middle) or monocyte-depleted (right) mice related to **b** at 11 days after tumour implantation. Red, T cells; green, KPP-eGFP. Scale bars, 2 mm. **e**, The proportion of immune-inflamed,

immune-excluded and immune-desert tumours related to **b**–**d**. **f**, The relative abundance of fibroblast subclusters determined by scRNA-seq analysis of STAMP tumour biopsies pooled by immune phenotype. **g**, The experimental design relating to **h**–**j**. DPT^{+} fibroblasts were ablated by tamoxifen and diphtheria toxin administration in Dpt-cre-ERT2 LSL-DTR mice before STAMP implantation of KPP-eGFP tumour arrays into mice reconstituted with $tdTomato^{+}$ T cells. $n = 5$ control mice, $n = 207$ tumours; $n = 7$ fibroblast-depleted mice, $n = 314$ tumours. **h**, Representative image of STAMP tumour arrays of control (top) or fibroblast-depleted (bottom) mice at 11 days after tumour implantation. Red, T cells; green, KPP-eGFP. Scale bars, 1 mm. **i**, The proportion of Immune-inflamed, immune-excluded and immune-desert tumours related to **g**. **j**, Flow-cytometry-based monocyte and dendritic cell profiling of tumours in fibroblast-depleted versus control non-depleted mice related to **g**. Data are mean ± s.e.m. Statistical analysis was performed using a two-tailed Mann–Whitney *U*-test.

Source Data

Further examination of the STAMP atlas revealed an enrichment of a unique subset of cancer-associated fibroblasts (CAFs) that we named ChemoCAFs in the inflamed and resolved tumours (Fig. 3f and Extended Data Fig. 8a). ChemoCAFs abundantly express chemokines, but lack expression of inflammatory cytokines, such as IL-1, IL-6, IL-11, LIF, CSF-2 or VEGFa, which are classically produced by inflammatory CAFs. ChemoCAFs also express dermatopontin (DPT), which is a known marker of $PI16^{+}$ fibroblasts, the primary fibroblast population at steady state in the skin[21] (Extended Data Fig. 8b–d) and the progenitors of CAF subsets in tumours[21]. We therefore used an inducible mouse model of DPT^{+} fibroblast depletion[21] to eliminate around 75% of the skin fibroblasts at the time of tumour implantation to inhibit the emergence of CAFs after STAMP tumour implantation (Fig. 3g, Extended Data Fig. 8e and Supplementary Data 4) and follow subsequent immune cell recruitment. In the absence of fibroblasts, STAMP microtumours exhibited a strong decrease in overall T cell recruitment and an increase in immune-desert tumours (Fig. 3h,i and Extended Data Fig. 8f,g), emphasizing a role for the fibroblasts of the tumour microenvironment in shaping the immunological infiltrate of STAMP tumours.

Given that fibroblast and monocyte/neutrophil depletion both produced an enrichment in the immune-desert phenotype, we hypothesized that there might be a mechanistic relationship between the two cell types. Importantly, 8 out of the 10 top marker genes defining the ChemoCAF population are chemoattractants for monocytes and neutrophils (CCL7, CCL2, CXCL1, CXCL5, CCL8, chemerin and CXCL12; Extended Data Fig. 8h). We used a cell–cell communication database, CellChat[22], which predicted a role for ChemoCAFs in recruiting monocytes and neutrophils (Extended Data Fig. 8i,j). Using flow cytometry, we confirmed experimentally that there was deficient myeloid and T cell recruitment in the CAF-depleted tumour microenvironment but not in a secondary lymphoid organ such as the spleen (Fig. 3j and Extended Data Fig. 8k,l).

Together, these results suggest that the early fibroblast niche signals to myeloid cells (including neutrophils and monocytes) that in turn promote T cell recruitment and infiltration. T cell recruitment into an inflamed tumour microenvironment supports their effector function and increases the probability that a given microtumour will be rejected.

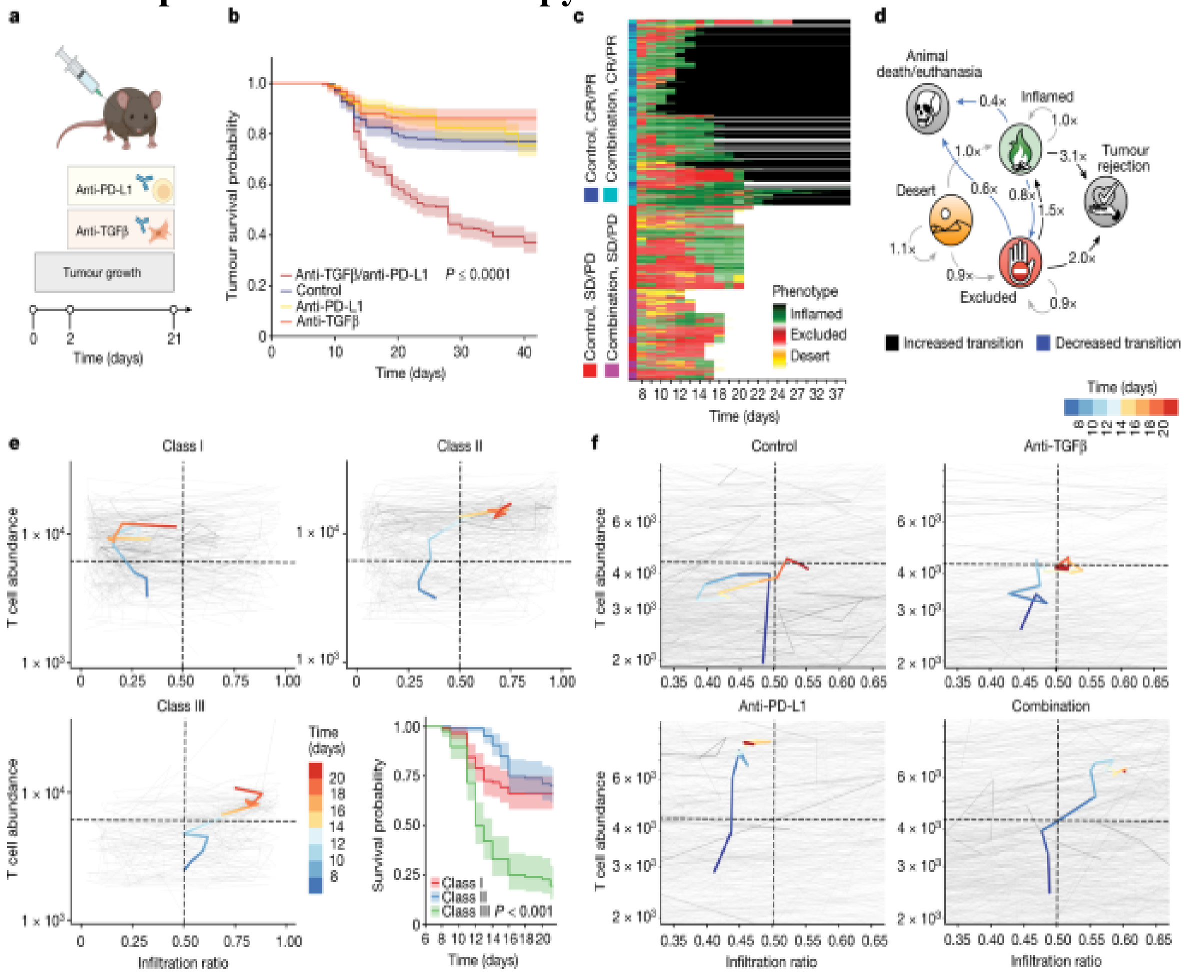
Early inflammation predicts rejection

Finally, we investigated how the spatial distribution of T cells predicts tumour progression or rejection during immunotherapy. Although tumours exhibit one or another immune phenotype, it is unclear whether these states are stable or dynamic over time. In humans, repeat biopsies before and after treatment can reveal alterations in immune phenotype (Extended Data Fig. 9a and Supplementary Table 7), but it is difficult to know whether these changes reflect an overall alteration in TIPs as opposed to pre-existing spatial heterogeneity[23,24]. By contrast, STAMP provides a unique opportunity to determine how immune phenotypes evolve spontaneously or after therapy.

As immune-excluded and immune-desert STAMP tumours exhibited an upregulation in the TGFβ pathway (Fig. 2b) and a combination of TGFβ and PD-L1 inhibition has previously shown anti-tumour efficacy in mice[11,25], we examined the effects of these agents in STAMP (Fig. 4a). In contrast to single-agent treatment, the combination therapy of anti-TGFβ and anti-PD-

L1 antibodies led to an improved overall response rate in KPP-eGFP STAMP tumours (Fig. 4b and Extended Data Fig. 9b–d).

Fig. 4: Early transition to an immune-inflamed phenotype predicts tumour response to immunotherapy.



a, The experimental design relating to the experiments shown in **b**–**f**. KPP-eGFP STAMP tumour arrays were implanted into RAG2-deficient mice reconstituted with tdTomato^{+} T cells and treated at day 2 after implantation with isotype control antibodies ($n = 554$ tumours, $n = 9$ mice), anti-TGFβ ($n = 287$ tumours, $n = 5$ mice), anti-PD-L1 ($n = 399$ tumours, $n = 6$ mice), or a combination of anti-PD-L1 and anti-TGFβ ($n = 642$ tumours, $n = 11$ mice). **b**, The survival probability of individual tumours of KPP-eGFP arrays related to **a**. Statistical analysis was performed using a log-rank test (referenced to the isotype control). **c**, Hierarchical clustering of individual tumour trajectories related to **a**, showing immune phenotypes over time for tumours treated with isotype control or a combination of anti-PD-L1 and

anti-TGFβ antibodies. Black, tumour resolved; white, mouse death/euthanasia; cyan, combination-treated responders (complete responders and partial responders (CR/PR)); magenta, combination-treated non-responders (stable disease and progressive disease (SD/PD)); blue, control responders (CR/PR); red, control non-responders (SD/PD). **d**, Markov chain showing the fold difference in the probabilities of transition between TIPs for combination anti-PD-L1/anti-TGFβ treatment versus the control condition. Bold indicates increased transition. Blue indicates decreased transition. '×' is the fold change. **e**, Unsupervised clustering of individual tumour immune trajectories highlighting changes in T cell abundance and infiltration ratio over time. $n = 6$ isotype-treated control mice. $n = 321$ tumours. The median immune trajectory for each of the three identified classes is shown in bold, and the colour scale indicates time. The survival probability of individual tumours grouped by immune-trajectory class is shown at the bottom right. Statistical analysis was performed using a log-rank test (referenced to class I tumours). **f**, Immune trajectories of individual tumours grouped by treatment related to **a**. Median immune trajectories are shown in bold, and the colour scale indicates time. In **b** and the bottom right of **e**, the centre line shows the Kaplan–Meier curve and the shaded area shows the 95% confidence interval.

Source Data

To better understand how T cell dynamics are associated with tumour rejection, we examined the spatial distribution of tdTomato$^+$ T cells over the lifetime of around 2,000 tumours using a high-content image analysis pipeline (Extended Data Fig. 9e–g). The immune history of individual tumours was plotted using a hierarchically clustered heat map to create representations of TIP, tumour growth rate and total T cell abundance over time in control and combination treated mice (Fig. 4c and Extended Data Figs. 9h and 10a–c). We found that TIPs are dynamic and can interconvert over time both in control and treated mice.

To quantify the effect of the combination therapy on the spatial distribution of T cells, we modelled phenotype transitions as a Markov chain to measure the probability of transition between the different states: immune-desert, immune-excluded, immune-inflamed, rejected tumour and mice death/euthanasia. We then compared the Markov chain for combination-

treated mice with the control (Fig. 4d, Supplementary Table 8 and Supplementary Data 5). We observed that the combination treatment increased the probability of transition from immune-excluded to immune-inflamed tumours and decreased the probability of reverse transition from immune-inflamed to immune-excluded tumours. Moreover, the probability of an immune-inflamed or an immune-excluded tumour becoming resolved was increased relative to control if treated with combination immunotherapy. These results highlight the overall dynamics of transitions between immune states promoted by combination immunotherapy.

To complement the Markov analysis and link the immunomodulatory effect of the treatment with its efficacy, we evaluated whether the history of an individual tumour might predict its eventual likelihood of rejection. We first defined the infiltration ratio for each tumour over time by measuring the T cell abundance in the tumour core relative to the total T cell abundance (Extended Data Fig. 9f). We then plotted the total T cell abundance against the T cell infiltration ratio for each individual tumour at every timepoint and performed unbiased clustering of the individual tumour trajectories. Unbiased clustering revealed three classes of tumour trajectories in the control condition (class I, II and III; Fig. 4e). Class I defines tumours that increase T cell abundance but progress as immune-excluded tumours; class II defines tumours transitioning first through an immune-excluded state before ultimately reaching a late inflamed phenotype; and class III defines tumours moving early to an immune-inflamed phenotype with a high T cell abundance and infiltration ratio. Notably, class III was the only trajectory associated with a strong enrichment in spontaneous tumour rejection, suggesting that considering the history of a TIP is essential to predict outcome.

To determine the treatment effect, we grouped individual tumour trajectories according to treatment, and represented the average trajectory for each treatment group (Fig. 4f and Extended Data Fig. 10d). Combination-treated tumours displayed a steady, strong increase in the total T cell abundance, and an inflamed T cell infiltration ratio similar to the class III trajectory. Tumours treated with anti-PD-L1 therapy were also marked by a strong increase in T cell abundance; however, they kept a very low infiltration ratio and persisted mostly as immune-excluded tumours similar to the class I

trajectory. By contrast, anti-TGFβ-treated tumours displayed the lowest T cell recruitment and infiltration ratio, similar to the control group.

Finally, we evaluated the trajectories of responders versus non-responder tumours in the combination therapy and showed that responding tumours followed the class III trajectory (Extended Data Fig. 10e) associated with an early infiltration of T cells in the tumour core, whereas the non-responders persisted with a low T cell abundance and low infiltration ratio. These results illustrate how combination immunotherapy pushes tumours towards the path of early inflammation leading to rejection, and they highlight the predictive value of TIP history.

Conclusion

Here we describe a preclinical model to characterize the spatiotemporal patterns of an antitumour immune response. STAMP revealed that genetically identical clonal tumours generate dynamic CD8 T cell infiltrations that are heterogenous between adjacent tumours. Longitudinal monitoring of the local immune response revealed that tumour rejection is associated with the early transition towards an immune-inflamed phenotype, characterized by heightened T cell cytotoxicity and a local decrease in TGFβ. We show that T cell infiltration is controlled in time and space by the tumour microenvironment (a summary is provided in Extended Data Fig. 10f) such as chemokine-producing CAFs and myeloid cells, both important players to recruit and support T cell function locally in inflamed tumours leading to tumour rejection. This study supports the role of an early inflamed TIP as a predictive biomarker of response to immunotherapy in patients by nurturing the function of newly expanded T cells. However, this immune phenotype is not stable over time, highlighting the importance of tracking the immune history of tumours in patients. Understanding the spontaneous or cancer immunotherapy-induced evolution of TIPs in the clinic will be an important consideration in predicting the likelihood of clinical response and providing biomarkers that may guide therapy in patients with cancer.

Methods

STAMP implantation

Experimental procedures were performed in 8–12-week-old male mice under anaesthesia (60–100 mg per kg ketamine and 5–10 mg per kg xylazine, intraperitoneal (i.p.) injection). Before microporation, the ear hair was removed using Veet depilatory cream, then rinsed with water, dried and the dorsal side of the ear was immobilized with double-sided tape to expose the ventral side. Microporation at 71 µm depth and 9% pore density was performed by applying the P.L.E.A.S.E. laser device (Pantec Biosolutions) to the ventral side of the ear using a custom program previously validated for the cell line. After the microporation process, 150 µl of a tumour cell suspension (EMT6-eGFP, CT26-RFP, 4T1-mTagBFP2, KPP-eGFP, KPP-mTagBFP2, KPP-eGFP-OVA, NSCLC-EGFP or B16F10) at 40–80 × 10^6 cells per ml diluted in PBS was applied to the ear covering the ventral side. Cells were incubated for 30 min then the excess cell suspension was removed, pores were covered with Matrigel (Corning) and incubated for 15 min until Matrigel polymerized. Microtumour growth becomes evident 5 to 12 days after tumour implantation, depending on the cell line. Once tumours were visibly observed, the mice were observed and tumours were measured at least twice per week. Monitoring frequency was increased if any additional adverse effects were observed, up to daily depending on the severity, according to the direction of the veterinary staff. The mice were euthanized immediately if the tumour area reached or exceeded 70% or the total ear surface, or if tumours fell outside the IACUC Guidelines for Tumours in Rodents. Moreover, if the ears were in a condition that compromises normal ear and bodily function as determined by the veterinary staff, the mice were euthanized. Any mice exhibiting other unexpected adverse effects such as severe hunching or severe lethargy or any moribund mice were euthanized immediately. Mice were weighed at least weekly and those mice losing >15% body weight were weighed daily. Animals that had lost 20% body weight were euthanized or brought to the attention of the veterinary staff. Mice with a body condition score of 2 or less out of 5 were euthanized.

Cell lines

B16F10, EMT6, CT26 and 4T1 mouse parental lines were sourced from the ATCC. KPP (PDAC) and NSCLC cells were derived from primary tumours of Genentech Cancer Immunology GEMM mice. Parental cell lines were further engineered in house using PiggyBac vectors to overexpress different types of fluorescent reporters (eGFP, RFP or mTagBFP2) or/and model antigens (ovalbumin or decamer peptide of tumour neoantigens, including the M86 neoantigen). Genentech built a centralized cell bank, gCELL, to support the needs of cell-based research within Genentech. gCELL is tasked to bank verified, quality-assured cell lines for distribution throughout Genentech. This provides a consistent source of cell lines for all levels of research to enable experimental reproducibility and access to baseline information such as morphology, growth conditions, RNA-seq and whole-exome sequencing derived from these lines. gCELL also provides an important mechanism to ensure that cell lines are used in accordance with all terms and conditions. All stocks are tested for mycoplasma before and after cells are cryopreserved. Two methods were used to avoid false-positive/negative results: Lonza Mycoalert and Stratagene Mycosensor. All of the cell lines tested negative for mycoplasma.

Mice

Mice were housed under specific-pathogen-free conditions at the Genentech animal facility. Mice were maintained in accordance with the Guide for the Care and Use of Laboratory Animals (National Research Council, 2011). Genentech is an AAALAC-accredited facility and all animal activities in this research study were conducted under protocols approved by the Genentech Institutional Animal Care and Use Committee (IACUC). Mice were housed in individually ventilated cages within animal rooms maintained under a 14 h–10 h light–dark cycle. Animal rooms were temperature and humidity controlled between 68 and 79 °F (20.0 to 26.1 °C) and from 30% to 70%, respectively, with 10 to 15 room air exchanges per hour. Male mice (aged 8–12 weeks) that appeared healthy and free of obvious abnormalities were used for the study. B6.Cg-Foxn1nu/J (000819), C57BL/6-Tg (CAG-EGFP)1Osb/J (003291) and C57BL/6J (000664) and B6(Cg)-Ifnar1 tm1.2Ees/J (028288) mice were purchased from Jackson Laboratories. B6.129S6-Rag2^{tm1Fwa} N12 (RAGN12), C.Cg/AnNTac-Foxn1^{nu} NE9

(BALBNU-M) and BALB/cAnNTac (BALB-M) mice were purchased from Taconic Biosciences.

CD4.cre.tg Rosa26.LSL.tdTomato.cki OT-I.TCR.tg (*OT-I*$^{-/-}$ and *OT-I*$^{+/+}$), DPT-IRES-Cre.ERT2. ki.B6N.1C9-1-H4-1_Rosa26.LSL.YFP.cki.B6J_Rosa26.LSL.DTR.cki and E8I.CD8A.IRES.GFP.Cre.tg Rosa26.LSL.tdTomato.cki mice were bred in house.

Subcutaneous tumours

For subcutaneous tumour inoculation, mice were injected subcutaneously with 0.5×10^6 KPP-eGFP cells in 100 µl of a 1:1 dilution of PBS and Matrigel (Corning). Once the tumours were visibly observed, the mice were observed and tumours were measured at least twice per week. The monitoring frequency was increased if any additional adverse effects were observed, up to daily depending on severity, according to the directions of the veterinary staff. Mice were euthanized immediately if the tumour volume exceeds 2,000 mm^3, or if tumours fell outside the IACUC Guidelines for Tumours in Rodents. Any mice exhibiting other unexpected adverse effects such as severe hunching or severe lethargy or any moribund mice were euthanized immediately. Mice were weighed at least weekly and those mice losing >15% body weight were weighed daily. Mice that lost 20% body weight were euthanized or brought to the attention of the veterinary staff. Mice with a body condition score of 2 or less out of 5 were euthanized.

In vivo cell depletion

Myeloid depletion

E8I.CD8A.IRES.GFP.Cre.tg Rosa26.LSL.tdTomato.cki mice were treated i.p. with 300 µg per mouse of anti-mouse Ly6C/G, anti-mouse Ly6C antibodies or isotypes controls from BioXcell (BE0075 and BE0203, respectively) for 3 consecutive days before KPP-eGFP STAMP tumour implantation. To maintain the cell depletion, the antibody treatment was sustained until the end of the study. Depletion was validated using flow

cytometry. Tumour growth and T cell infiltration was monitored daily by GFP/tdTomato imaging.

CD8 and CD4 T cell depletion

Mice were treated with 100 µg per mouse i.p. of anti-mouse CD8b (BE0223) and anti-mouse CD4 (BP-0003-3) or isotype control: (1) C57BL/6J mice were CD8 and CD4 depleted for 3 consecutive days before KPP-eGFP STAMP tumour implantation. Tumour growth was followed daily by direct imaging. (2) DPT.IRES.Cre.ERT2.ki. B6N.1C9-1-H4-1_ Rosa26.LSL.YFP.cki. B6J_Rosa26.LSL.DTR.cki mice were injected with anti-CD8 and anti-CD4 antibodies for 3 consecutive days. After 14 days (when antibody was cleared) the DTP-diphtheria-toxin-induced depletion experiment started. All CD4 and CD8 depletions were confirmed using flow cytometry.

Fibroblast depletion

As a DPT expression control, DPT.IRES.Cre.ERT2.ki. B6N.1C9-1-H4-1_Rosa26.LSL.YFP.cki. B6J_Rosa26.LSL.DTR.cki mice, previously CD4/CD8 depleted, were injected daily with 20 µg per mouse i.p. of tamoxifen. After DPT–YFP expression was established by flow cytometry, mice were treated with diphtheria toxin every other day starting at 4 days before KPP-eGFP STAMP tumour implantation and tdTomato^{+} T cell reconstitution. Depletion was validated using flow cytometry. Tumour growth and T cell infiltration was monitored daily by GFP/tdTomato imaging.

Adoptive transfer of antigen-specific T cells

B6.129S6-Rag2tm1Fwa N12 mice were adoptively transferred with 4×10^6 *OT-I*$^{-/-}$ tdTomato^{+} T cells or 4×10^6 *OT-I*$^{+/+}$ tdTomato^{+} T cells after KPP-EGFP-OVA expressing cells were implanted. Tumour growth and T cell infiltration was monitored daily by GFP/tdTomato imaging

Cancer immunotherapy treatment

Antibody treatment

Mice were implanted with STAMP microtumours as described above. Mice were distributed into treatment groups to exclude cage effects and, when possible, to account for differential initial tumour growth. Treatment was initiated 1 day after tumour implantation for immunodeficient mice or 7 days after tumour implantation for immunocompetent mice. Mice were treated every other day with isotype control antibodies (anti-gp120; mouse IgG, 3E5, 10 mg per kg), anti–PD-L1 (mouse IgG1, 6E11, 10 mg per kg first dose followed by 5 mg per kg thereafter), anti-TGFβ (mouse IgG1, 1D11, 10 mg per kg) or a combination of anti-PD-L1 with anti-TGFβ. Beginning 5 days after treatment initiation, mice were imaged daily using the M205FA stereoscope with a ×1.0 PlanApo lens (10450028; Leica Microsystems) and the ORCAII Digital CCD (Hamamatsu Photonics) to monitor tumour growth and T cell progression. For selected experiments, immunodeficient mice were injected intravenously with 4×10^6 isolated tdTomato$^+$ T cells from CD4.cre.tg_Rosa26.LSL.tdTomato.cki_OT-I.TCR.tg mice as described in the figure legends. Individual microtumours were biopsied individually or pooled by phenotypes.

Tumour antigen vaccination

The KPP-GFP clonal cell line was CRISPR-edited to ablate GFP expression, and then transfected with a PiggyBac vector to express a decamer peptide of tumour neoantigens, including the M86 neoantigen, and cytoplasmic mTagBFP2 (EF1A_MC38-Deca16-cBFP). M86 RNA-lipoplex (RNA-LPX) vaccines were assembled from M86-coding RNAs synthesized by Genentech with liposomes consisting of DOTMA and DOPE at a charge ratio (+):(−) of 1.3:2.0 as described previously[26]. CD4.cre.tg Rosa26.LSL.tdTomato.cki T cell donor mice ($n = 2$) were intravenously vaccinated with M86 RNA-lipoplex 1, 2 and 3 weeks before T cell isolation using the EasySep Mouse T Cell Isolation Kit (StemCell Technologies) from the spleen. T cells were used for adoptive transfer into RAG2-deficient mice or in vitro stimulation assays. To validate RNA-LPX-induced antigen-specific T cells, 200,000 splenocytes from vaccinated ($n = 2$) or naive ($n = 2$) mice were co-cultured with 40,000 tumour cells expressing (+) or not (−) the tumour neoantigens. As a positive control, M86 peptide was added to the

negative tumour cells. Antigen-specific T cell stimulation was measured by flow cytometry as PD-1^+ percentage of live, $CD45^+CD90.2^+CD8^+CD44^+SlamF7^+$ cells. STAMP tumour assays were performed as described in RAG2-deficient mice with 4×10^6 $tdTomato^+$ $CD3^+$ T cells isolated as described below from vaccinated or naive mice.

Metastasis mouse model and imaging of whole-mount tissue

RAG2-deficient mice were injected intravenously with 0.1×10^6 KPP-eGFP cells and 4×10^6 $tdTomato^+$ $CD3^+$ T cells isolated as described below. Then, 8 days after intravenous injection, the mice were euthanized, and 20 ml of cold PBS/heparin 5 IU ml^{-1} solution was perfused directly into the right ventricle using a 27-gauge needle. Lungs were isolated by dissection and tissues were fixed using 4% paraformaldehyde in PBS[23]. Tissue clearing was performed using the FluoClearBABB approach[27] and whole-mount images were then acquired using a SP8 microscope equipped with a white-light laser and the HCX APO L ×20/0.95 NA IMM lens (Leica Microsystems). Imaging data were analysed on a workstation (Thinkmate) using Imaris software (Bitplane).

T cell isolation

T cells were isolated from C57BL/6J or CD4.cre.tg Rosa26.LSL.tdTomato.cki OT-I.TCR.tg ($OT\text{-}I^{-/-}$ and $OT\text{-}I^{+/+}$) mice. Spleens were collected and dissociated with the end of a plunger from a 1 ml syringe into 10 ml of PBS before filtration through a 70 μm cell strainer. T cells were isolated for intravenous injection using the EasySep Mouse T Cell Isolation Kit (StemCell Technologies). A total of 4×10^6 isolated T cells was injected intravenously per mouse.

To generate cytotoxic T cells from OT-I mice, splenocytes were isolated and stimulated with 10 nM OVA257-264 peptide (AnaSpec) in complete medium containing Gibco RPMI 1640 (Thermo Fisher Scientific) with 10% Gibco fetal calf serum (Thermo Fisher Scientific), 2 mM l-glutamine, 50 IU ml^{-1} Gibco penicillin–streptomycin (Thermo Fisher Scientific) and 50 μM β-mercaptoethanol. After 3 days of stimulation, cells were resuspended in

complete medium with 10 IU ml^{-1} recombinant human IL-2 (rHIL-2). Cytotoxic T cells were kept at a density of 0.5×10^6 cells per ml and fresh complete medium with rHIL-2 was added every 48 h. Cytotoxic T cells were used between 6 and 8 days after primary in vitro stimulation.

Flow cytometry

Ear tissue was isolated using a 1 mm Miltex sterile disposable biopsy punch with plunger (Integra Biosciences) from mice bearing STAMP microtumours or mock-implanted control mice at either 8 or 18 days after tumour implantation. Tumours were digested in 500 µl (STAMP microtumours) or 3,000 ml (subcutaneous) of PBS containing 0.1 mg ml^{-1} DNase I (Roche) and collagenase D at 1 mg ml^{-1} (Roche) for 30 min at 37 °C to obtain a single-cell suspension.

For surface staining, cells from digested tumours were incubated with Fc block (5 µg ml^{-1}, BD Biosciences, 2.4G2) and stained with antibody mix for 30 min and Viability Dye eFluor 780 (eBioscience).

Antibodies were used at 1:200 dilution. Anti-mouse CD19 (B4), anti-mouse I-A/I-E (M5/114.15.2), anti-mouse F4/80 (BM8), anti-mouse/human CD11b (M1/70), anti-mouse Ly-6G (1A8), anti-mouse Ly-6C (HK1.4), anti-mouse CD69 (H1.2F3), anti-mouse CD25 (PC61), anti-mouse CD4 (RM4-5), anti-mouse CD62L (MEL-14), and anti-mouse/human CD44 (IM7) antibodies were purchased from BioLegend. Anti-mouse CD45 (30:F11), anti-mouse CD86 (GL1), CD11c anti-mouse (N418), and anti-mouse CD8a (53-6.7) antibodies were purchased from Thermo Fisher Scientific. Anti-mouse CD3 (17A2) antibodies were purchased from BD Biosciences.

Live $CD45^+$ singlet cell subsets were gated as follow: MHC class $II^+CD11c^+F4/80^-$ (dendritic cells) or $CD103^+$ or $CD86^+$ (activated dendritic cells); $CD11c^-CD11b^+F4/80^+$ (macrophages); $CD11b^+Ly6G^+Ly6C^{int}$ (neutrophils); $CD11b^+Ly6G^{low}Ly6C^+$ (monocytes); $CD3^+$ T cells were divided as $CD3^+CD4^+$ T cells, $CD3^+CD8^+$ T cells, $CD3^+CD69^+$ activated/resident T cells, $CD3^+CD44^+CD62L^-$ effector/effector memory T cells or $CD3^+CD44^-CD62L^+$ naive T cells.

Flow cytometry data were collected using the BD LSRFortessa cell analyser (BD Biosciences) and analysed using FlowJo Software (v.10.2; FlowJo).

STAMP microtumour sample preparation for bulk RNA-seq of tumour biopsies

FOXN1-deficient nude mice with adoptively transferred tdTomato^{+}CD3^{+} T cells bearing KPP-eGFP STAMP microtumours were biopsied 8 days after tumour implantation using a 1 mm Miltex sterile disposable biopsy punch with plunger (Integra Biosciences). Each tissue biopsy was transferred into a separate 1.5 ml tube (Eppendorf) containing 0.25 ml Invitrogen TRIzol Reagent (Thermo Fisher Scientific). Biopsies were incubated in TRIzol for 5 min with intermittent vortexing. A total of 50 µl chloroform (MilliporeSigma) was added to the homogenate, vortexed for 20 s and incubated at 20–25 °C for 2–3 min. To accelerate phase separation, the samples were centrifuged at 10,000*g* for 18 min at 4 °C. The aqueous (top) phase was removed by aspiration and transferred to a clean 1.5 ml tube (Eppendorf). A volume of 100% RNase-free ethanol (MilliporeSigma) equal to the volume of the aqueous layer was added, and the RNA was further isolated using the RNeasy Micro Kit (Qiagen). Alternatively, an individual tumour biopsy was immersed in RNA later and the RNA was further extracted using the RNeasy Mini Kit (Qiagen).

Sample preparation of sorted STAMP microtumours for scRNA-seq

FOXN1-deficient nude mice with adoptively transferred tdTomato^{+} CD3^{+} T cells bearing KPP-eGFP STAMP microtumours were biopsied 8 days after tumour implantation using a 1 mm Miltex sterile disposable biopsy punch with plunger (Integra Biosciences). Then, 3–6 pooled tissue biopsies were moved into a precooled 1.5 ml tube (Eppendorf) containing 300 µl digestion cocktail consisting of Gibco RPMI 1640 (Thermo Fisher Scientific), 0.1% Gibco fetal calf serum (Thermo Fisher Scientific), 0.1 mg ml^{-1} Liberase (Roche), 0.1 mg ml^{-1} DNase I (Roche) and 32 µM Gibco actinomycin D (Thermo Fisher Scientific). Tissues were incubated for 30 min at 37 °C and 950 rpm on a Thermoblock (Eppendorf) and mechanically dissociated every

10 min with a pipette. To quench the digestion, the cell suspension was filtered through a 40 µm mesh into a precooled fluorescence-activated cell sorting (FACS) filter tube containing quenching buffer of Gibco fetal calf serum (Thermo Fisher Scientific) with 32 µM Gibco actinomycin D (Thermo Fisher Scientific). The cell suspension was centrifuged at 350*g* for 8 min and resuspended in 400 µl Gibco RPMI 1640 (Thermo Fisher Scientific) with 5 µM Calcein Blue (Invitrogen) and a 1:200 dilution of the Molecular Probes Fixable Live/Dead Near-IR Dead Cell Stain Kit (Thermo Fisher Scientific). FACS analysis was performed to isolate Calcein-Blue-positive and L/D NearIR-negative cells into a precooled 1.5 ml Eppendorf tube containing 750 µl collection buffer consisting of Gibco RPMI 1640 (Thermo Fisher Scientific), 10% Gibco fetal calf serum (Thermo Fisher Scientific), 32 µM Gibco actinomycin D (Thermo Fisher Scientific). The cell number and viability were determined on the Vi-Cell XR cell viability analyser (Beckman Coulter) and scRNA-seq library preparation was performed using the dual single-cell mouse kit 5′/TCR according to the manufacturer's instructions (10x Genomics).

3D in vitro tumouroid cultures/co-cultures

KPP mouse pancreatic cancer cells expressing human HER2 and cytosolic GCaMP6 were suspended in three-dimensional collagen matrices as described previously[28]. In brief, a solution of rat-tail collagen I (MilliporeSigma) was brought to a neutral pH on ice and mixed with KPP.hHer2.GCaMP6 cells to a final concentration of 2 mg ml^{-1} collagen and 1.0×10^4 cells. Then, 150 µl of this suspension was added to individual wells of an eight-chambered cover glass (Cellvis). The chambers were incubated at 20–25 °C for 15 min then incubated at 37 °C with 5% CO_2 for an additional 15 min. After incubation, 500 µl of complete medium containing Gibco RPMI 1640 (Thermo Fisher Scientific) with 10% Gibco fetal calf serum (Thermo Fisher Scientific), 2 mM l-glutamine and 50 IU ml^{-1} Gibco penicillin–streptomycin (Thermo Fisher Scientific) was carefully added to each well. Cells were allowed to grow in collagen matrices for 5 days before imaging.

3D imaging of OT-I T cells interacting with KPP.hHer2.GCaMP6 tumouroids in collagen matrices was performed on the TiE microscope

(Nikon) with the CSU-X1 Spinning Disk (Yokogawa Electric) and Prime sCMOS camera (Photometrics). The medium in the eight-chamber imaging slides containing tumour cell collagen matrices was replaced with Gibco RPMI 1640 with no phenol red (Thermo Fisher Scientific) with 10% Gibco fetal calf serum (Thermo Fisher Scientific), 2 mM l-glutamine and 50 IU ml^{-1} Gibco penicillin–streptomycin (Thermo Fisher Scientific) with 3 µM propidium iodide (Thermo Fisher Scientific). OT-I cells were prelabelled with Celltrace FarRed (Thermo Fisher Scientific) according to the manufacturer's protocol. FarRed-labelled OT-I cells were added to each chamber containing collagen-suspended KPP.hHer2.GCaMP6 tumouroids and allowed to infiltrate for 2 h before imaging. At the time of imaging, T-cell-dependent bispecific antibody (anti-hHER2::anti-CD3e) was added to the chambers at a final concentration of 500 nM to induce cytotoxic T cell recognition of hHER2-expressing cancer cells. In the control conditions, no antibody was added.

STAMP microtumour correlative imaging of Ca^{2+} influx and T cell infiltration

RAG2-deficient mice with adoptively transferred $tdTomato^{+}$ $CD3^{+}$ T cells bearing mTagBFP2 and GCaMP6-expressing KPP STAMP microtumours were anaesthetized by isoflurane inhalation to effect and imaged daily from 4 days after tumour implantation to 15 days after tumour implantation at the same ear regions. Epifluorescence time-lapse microscopy image series were acquired daily at the same ear regions with the ×1.0 Leica PlanApo objective (Leica 10450028) on the Leica M205 FA epifluorescence stereomicroscope every minute for 60–70 min. Image analysis was performed using Imaris (Bitplane). Time-lapse image series of individual tumours at 8 days after tumour implantation were semi-automatically segmented and analysed for Ca^{2+} influx between timepoints. Furthermore, tumour sizes, T cell abundances and T cell infiltrations were analysed. Time-lapse image sequences of individual tumours at 13 days after tumour implantation were semi-automatically segmented and were analysed for tumour size to determine the tumour growth from day 8 to day 13.

STAMP microtumour correlative imaging of Ca^{2+} influx and T cell infiltration after TDB administration

FOXN1-deficient nude mice with adoptively transferred $tdTomato^{+}CD3^{+}$ T cells bearing mTagBFP2 and GCaMP6-expressing KPP STAMP microtumours were anaesthetized by isoflurane inhalation to effect and imaged 12 days after tumour implantation. Image series were acquired every 90 s for 45 min using a two-photon laser-scanning microscope (Ultima In Vivo Multiphoton Microscopy System, Bruker Technologies) with alternating excitation from dual Ti:sapphire lasers (MaiTai DeepSee, Spectra Physics) tuned to 830 nm and 980 nm, and a ×16/0.8 NA immersion objective lens (Nikon). Thereafter, T-cell-dependent bispecific antibodies (anti-hHER2::anti-CD3e) were administered intravenously (6 mg per kg) and multiphoton time-lapse microscopy image acquisition was continued at the same region. Time-lapse image series of individual tumours were semi-automatically segmented with Imaris (Bitplane) and analysed for Ca^{2+} influx between timepoints.

Immunophenotypes in the imCORE Paired Biopsy trial

Tumour biopsies were obtained from patients enrolled in the imCORE[29] Paired Biopsy trial (NCT03333655) between January 2018 and March 2020. This study is an ongoing, open-label, multicentre trial initiated in February 2018 and conducted globally including study centres in the USA, France and Spain. Adult patients with metastatic cancer or haematological malignancies who demonstrated clinical benefit on cancer immunotherapy and had a tumour biopsy both at baseline (pre-treatment/archival) and at progression were eligible for inclusion. Cancer immunotherapy included marketed agents (including those targeting CTLA-4, PD-L1 or PD-1) or those administered through participation in a Roche/Genentech CPI clinical trial. Patients with the best overall response (on the basis of Response Evaluation Criteria in Solid Tumours v.1.1) of complete response, partial response or stable disease after >6 months (or >3 months if enrolled under an earlier protocol version) were eligible. PanCK/CD8 dual staining was performed on histological sections from baseline and progression formalin-fixed paraffin-

embedded tumour samples. Immune phenotypes were determined by a pathologist (Histogenex) using defined criteria[30].

Clinical trials

Clinical trials registrations were as follows: IMvigor210 (NCT02951767/NCT02108652), ICON7 (NCT00483782) and imCORE trial (NCT03333655). For the ICON7 and IMvigor210 studies, the full protocols are available at https://clinicaltrials.gov/. For the imCORE trial, the protocol is available on request (www.roche.com/about_roche/roche_worldwide.htm, +1 888-662-6728, global-roche-genentech-trials@gene.com). For the IMvigor210 trial, the study was approved by the independent review board at each participating site and was performed in full conformance of the provisions of the Declaration of Helsinki and Good Clinical Practice Guidelines. Approval from the Institutional Ethics Committee or the Institutional Review Board was obtained before the study start and was documented in a letter to the investigator specifying the date on which the committee met and granted the approval. The ICON7 protocol was compliant with good clinical practice guidelines and the Declaration of Helsinki. Approval by ethics committees was obtained at each clinical site, nationally or both. For the imCORE trial, the study protocol was approved at enrolling institutions and by local ethics committees (Sarah Cannon Research Institute, WIRB; IUCT Oncopole Toulouse, France; Clinica Universidad di Navarra, Spain).

The ICON7 and IMvigor210 trials have been previously published[11,13]. For the imCORE trial, the patients were recruited by participating institutions if eligibility criteria (including clinical benefit from checkpoint inhibition and biopsies were available before and after treatment from the same tissue) were met. No knowledge of immunophenotype was known at the time of recruitment, therefore limiting potential bias. Clinical characteristics of imCORE patients are summarized in the Supplementary Table 7.

All patients have provided written informed consent.

Image analysis

Ca^{2+} influx index in vivo epifluorescence microscopy

An isosurface is created that matches the tumour-associated mTagBFP2 fluorescence of individual STAMP microtumours. The sum of mTagBFP2 and GCaMP6 fluorescence pixel intensities is calculated for each channel for the tumour isosurface for each timepoint. The absolute delta of the sum fluorescence intensities between consecutive timepoints is calculated, averaged for each fluorophore and normalized to the respective mean fluorescence intensity (MFI). The Ca^{2+} influx index is the result of dividing the normalized average delta sum of GCaMP6 intensities by the normalized average delta sum of mTagBFP2 intensities.

T cell abundance

An isosurface is created that matches the tumour-associated mTagBFP2 fluorescence of individual STAMP microtumours. The MFIs for the T cell-associated tdTomato fluorescence are determined for the tumour isosurface for each timepoint. The T cell abundance index is the result of calculating the median of the tdTomato MFIs across all timepoints.

T cell infiltration index

An isosurface is created that matches the tumour-associated mTagBFP2 fluorescence of individual STAMP microtumours. Using the tumour isosurface, two new regions are defined: the tumour centre (central 50% of tumour isosurface) and the tumour periphery (the area surrounding the tumour that is up to 50 µm distance from tumour border). The MFIs for the T-cell-associated tdTomato fluorescence are determined for the tumour centre and the tumour periphery for each timepoint, and the median of MFIs across all timepoints is calculated. The T cell infiltration index is the result of the ratio of those medians (centre/periphery).

Ca^{2+} influx index in vivo two-photon microscopy

An isosurface is created that matches the tumour-associated mTagBFP2 fluorescence of individual STAMP microtumours. The sum of mTagBFP2

and GCaMP6 fluorescence pixel intensities are calculated for the tumour isosurface for each timepoint. The absolute delta of the sum fluorescence intensities between consecutive timepoints is calculated, averaged for each fluorophore and normalized to the respective mean MFI across all timepoints. The normalized average delta sum of GCaMp6 intensities is divided by the normalized average delta sum of mTagBFP2 intensities (value1). Also, the average s.d. of mTagBFP2 and GCaMP6 fluorescence of every pixel of the tumour isosurface is calculated across the time series. The average s.d. of GCaMP6 is divided by the average s.d. of mTagBFP2 (value2). The Ca^{2+} influx index (two-photon) is the result of multiplying value1 and value2.

Ca^{2+} influx index in vitro spinning-disk confocal microscopy

An isosurface is created that matches the tumour-associated GCaMP6 background fluorescence of tumour cells. The median of GCaMP6 MFIs is calculated for the tumour cell isosurfaces for each timepoint. The absolute delta of the median MFIs between consecutive timepoints is calculated and averaged. The Ca^{2+} influx index (spinning-disk confocal microscopy) is the result of dividing the average delta median GCaMP6 MFIs by the mean GCaMP6 MFI across all timepoints.

PI influx index in vitro spinning disk confocal microscopy

An isosurface is created that matches the tumour-associated GCaMP6 background fluorescence of tumour cells. The median of propidium iodide MFIs is calculated for the tumour cell isosurfaces for each timepoint. The absolute delta of the median MFIs between consecutive timepoints is calculated and averaged. The propidium iodide influx index is the result of dividing the average delta median propidium iodide MFIs by the mean propidium iodide MFI across all timepoints.

U-net model training

Images of tumour fluorescence were binned and resized to 512 × 512 px using custom FIJI scripts. Binary (two-class) masks were manually

generated with 1 = tumour, 0 = background. A TensorFlow U-net model adapted from https://github.com/zhixuhao/unet was trained on a dataset of 595 paired images with masks (70% training and 30% validation) for 7 epochs until the model began to overfit as indicated by the training accuracy exceeding the validation accuracy without improving loss.

Image segmentation and tracking

Images of tumour fluorescence were binned and resized to 512 × 512 px using custom FIJI scripts. Initial segmentation guesses were generated by applying the trained U-net using TensorFlow. Custom FIJI scripts were used to un-bin tumour segmentation to restore the original size and resolution, enable manual review and editing of all tumour segmentation masks, and manually track tumours through multiple timepoints. If a tumour was no longer detectable during the course of an experiment, it was designated a complete responder. If a tumour decreased from its maximum size by 20% or more, it was designated a partial responder. The remaining tumours were designated as stable disease and progressing disease.

T cell quantification

Custom FIJI scripts were used to identify the centroid and Feret diameter of each tumour region of interest and determine the median radial fluorescence profile of the T cell fluorescence channel. Custom Python scripts were used to determine the overall median T cell fluorescence intensity and categorize radial fluorescence profiles as desert, excluded or inflamed. Tumours were classified as excluded or inflamed using a ratiometric cut-off. If the radial profile within the inner 25% of the tumour was consistently greater than 60% of the maximum fluorescence for that tumour, it was designated as inflamed. If the radial profile within the inner 25% of the tumour was consistently less than 40% of the maximum fluorescence for that tumour, it was designated as excluded. A tumour was designated a desert if the individual tumour's median T cell fluorescence intensity was less than the 25 percentile of median T cell fluorescence for all tumours measured on the first imaging day and the radial profile did not indicate an excluded pattern as described above. If a tumour showed an excluded profile based on the ratiometric criteria, but the T cell intensity at the core of the tumour (inner

25%) was greater than the median T cell intensity for all inflamed tumours, it was reclassified as inflamed. If a tumour did not meet the above ratiometric cut-offs, the phenotype determined the previous day was propagated forward until the next definitive classification.

Clustering and Markov analysis

Custom Python scripts were used to assign an integer value to the T cell phenotype classification with 1 = desert, 2 = excluded, 3 = inflamed. If the mouse was euthanized, the remaining timepoints were assigned a value of 0. If the tumour resolved, the remaining timepoints were assigned a value of 5. Tumour trajectories were ordered by hierarchical clustering of tumour phenotype lists. Subsequent heat maps for tumour area, median T cell infiltration and tumour growth were ordered according to this phenotype clustering. Transition state matrices for Markov analysis were generated using custom Python scripts. To assess significance, phenotype states were randomized ten times using Python random.shuffle() and new transition state matrices were calculated.

T cell trajectory analysis

Custom python scripts were used to calculate the infiltration ratio where the numerator is the median intensity at the core of the tumour (inner 25%), and the denominator represents the maximum value of the median radial profile. These values were plotted relative to the total T cell abundance and compared between tumours across timepoints.

Whole-exome sequencing analyses

To rule out the possibility that TIPs could be explained by pre-existing major subclones within the KPP-eGFP cell line, the tumour cell line genetic heterogeneity was assessed using whole-exome sequencing before implantation and 3 weeks after implantation. A spleen sample from the wild-type mouse strain used in the experiments was used as the reference matched normal genome to perform variant calling in tumour cells. The analysis after tumour implantation was restricted to the tumour variants detected in the KPP-eGFP cell line to alleviate the effect of any potential genetic variability

within mice captured from the different TME components of the different biopsies.

Exome capture libraries were sequenced on the HiSeq 2500 (Illumina) system to generate 2 × 75 bp paired-end data; from the sequenced reads, variants were called using the following workflow. Sequencing reads were mapped to the UCSC mouse genome (GRCm38) using BWA software[31] set to the default parameters. Local realignment, duplicate marking and raw variant calling were performed according to GATK best practices[32]. Somatic variant calling on tumour and its matched normal BAM file was performed using Strelka[33].

The resulting variants can be interpreted as a description of the genetic differences in the exome present in the KPP-eGFP cell line with regard to the host mouse genome. This measurement coarse grains over the clonal structure of the population; nevertheless, the inspection of the resulting variant allele frequency histogram can inform of potential subclones. For the preimplantation cell line samples, the variant allele frequency was summarized with a histogram. No major subclones were observed in these representations; however, as expected, there was evidence of minor variants. These are potentially neutral minor subclones, which are unavoidable during culturing practices such as an expansion phase after single-cell subcloning. To demonstrate that this low-frequency tail of minor subclones conformed to a neutral accumulation of mutants in the population, the left tail of the histograms was fit to a neutral model using the R package neutralitytestr (https://github.com/marcjwilliams1/neutralitytestr). For the tumour samples after implantation, some potential variants were detected as they were consistently uniquely present in either the immune-inflamed, the immune-excluded or the immune-desert tumours. To exclude having missed the variants due to the algorithm settings and the sequencing depth of each sample, the problematic variants were manually curated by inspecting the reads mapping to the regions in which the variants were detected on the remaining samples using the Integrative Genome Viewer. We therefore confirmed that, out of 48 potentially explanatory variants across immune phenotypes, all but one was detected on the other samples. The missing variant was a very-low-frequency variant that could be explained by not reaching the detection limit in the other tumour samples.

STAMP microtumour bulk RNA-seq analysis

Each RNA-seq experiment was analysed using HTSeqGenie pipeline in BioConductor[34] as follows: first, reads with low nucleotide qualities (70% of bases with quality < 23) or matches to rRNA and adapter sequences were removed. The remaining reads were aligned to the mouse reference genome GRCm38.p5 using GSNAP (v.2013-10-10-v2), allowing a maximum of two mismatches per 75 base sequence (parameters: '-M2 -n 10 -B 2 -i 1 -N 1 -w 200000 -E 1 --pairmax-rna = 200000 --clip-overlap'). Transcript annotation was based on the Gencode genes database[35]. To quantify gene expression levels, the number of reads mapping unambiguously to the exons of each gene was calculated.

The resulting count matrix was analysed in R (v.4.0.5; 31 March 2021) using the edgeR package (v.3.32.1). The count matrix was filtered to remove low-expressed genes by keeping the features with at least 0.2 counts per million (cpm) in more than the minimum number of samples of the experimental group given the design factor being analysed. The resulting filtered count matrix was then log-transformed and TMM-normalized with edgeR::cpm(log=T) and edgeR::calcNormFactors(method = "TMM") to perform exploratory analyses. The matrix was further filtered to the most variable genes selected using projection score[36] to focus on the major contributors to the variance in transcriptional state. This matrix was then scaled using base:scale() for posterior exploration. A heat map of the resulting data matrix annotated by Gene Ontology (GO) terms was constructed for preliminary interpretation by first clustering the genes and then running enrichment analysis using clusterProfiler (v.3.18.1) to select the most significant GO terms associated with said clusters (our script also allowed us to use WikiPathways and KEGG). To understand the general clustering of the samples in transcriptional state, dimensionality reduction of the scaled matrix using principal component analysis was performed using PCAtools (https://github.com/kevinblighe/PCAtools; v.2.5.15).

Differential expression analyses were conducted using limma (v.3.46.0)[37]. To detect signature differences among immune phenotypes, our contrasts compared the immune phenotype at hand with the average of the other two. Volcano plots were constructed using the package Enhanced-Volcano

(https://github.com/kevinblighe/EnhancedVolcano; v.1.8.0). Gene set enrichment analysis was performed on the log-transformed fold change given by the differential expression contrasts using the GSEA function from ClusterProfiler on the Hallmark Gene Set Collection[38]. Pathways were considered to be significant if their false-discovery-rate-adjusted *P* value was less than 0.2. The same analysis was performed on the clinical trial RNA-seq data. Heat maps of the normalized enrichment score were constructed for those significant pathways. For particular pathways that were further investigated in detail, samples were scored using the first principal component of the expression matrix of the genes composed the signature multiplied by the sign of the correlation of the component with the sample average expression of all the signature genes. For the time-series experiment, the type I interferon signature was extracted from the Hallmark Gene Set Collection.

scRNA-seq analysis of STAMP microtumours

scRNA-seq fastq files were processed using CellRanger count (CellRanger v.4.0.2 from 10x Genomics) using a custom reference based on the mouse genome GRCm38 with GENCODE[35] gene models together with the sequences for the transgenes used in our experiments (tdTomatoand GFP). Similarly, TCR sequencing data were processed using CellRanger vdj.

Analyses of the count matrices were conducted in R (v.4.0.5; 31 March 2021) using Seurat[39] (v.4.0.4). Only high-quality cells were retained for the posterior analysis; more concretely, we retained the cells with more than 300 hundred genes detected, more than 1,000 unique molecular identifiers and less than 10% mitochondrial reads. To simplify our analyses, tdTomato$^+$ cells (adoptive T cells) and GFP$^+$ cells (KPP tumour cells) were separated and clustered independently. The remaining cells were clustered to construct an atlas for stamp tumours. In both cases, the clustering and identification of cell populations proceeded following Seurat's SCTransform pipeline. First, data were normalized using Seurat::SCTransform() with cell cycle regression and batch correction. The number of principal components retained for clustering was then calculated using the talus plot[40]. With these retained components, we then computed a UMAP embedding and the neighbours for posterior clustering. Several clustering resolutions were

calculated and a directed tree was constructed reflecting the hierarchical relationships of the new clusters after increasing the resolution[41]. A resolution was considered to be optimal if it did not break the hierarchical structure of the said tree. The main clusters were identified by means of expression markers of known biology. Then, for each major cell cluster, a second run of clustering was performed by iterating the aforementioned pipeline. Notably, this step enabled us to identify heterogeneity within our cell populations and further remove low-quality cells. Markers for each cluster were identified using the Seurat::FindAllMarkers() method with the default parameters, comparing all cells in a particular cluster to the rest of cells and accessing significantly differential gene expression using Wilcoxon's rank-sum tests and Benjamini–Hochberg correction for multiple testing.

We used Seurat's plotting functionalities for most plots. Maker heat maps were generated using the package ComplexHeatmap using results from the Seurat::AverageExpression() function as the input after scaling to relative expression per gene using the *z*-score. Differential expression analyses were performed using Seurat::FindMarkers(), with batch as a latent variable and the negative binomial test. We reported the significant gene log-transformed fold change values as a *z*-score-scaled matrix using ComplexHeatmap. Clonotype analysis and integration with Seurat were performed using the scRepertoire package (v.1.0.0). Clonotypes were called according to their TCR amino acid sequence. Gene set enrichment analysis of the top seven dominant clonotypes was performed on the log-transformed fold change values from Seurat::FindMarkers(), comparing the inflamed to the excluded phenotypes using gseGO() function of the ClusterProfiler R Package on the CC Ontology collection.

To demonstrate the lack of deterministic correlation between the transcriptional heterogeneity of the KPP-eGFP cell line in vitro and the immune phenotypes in vivo, paired samples of the cell line before and after implantation were single-cell sequenced. Tumours were cell-hashed with antibodies and were then pooled before sequencing. The samples were demultiplexed after centred log ratio transformation normalization on the hashing antibody counts using Seurat::HTODemux() and doublets and negative cells were removed. Both the cell-hashed samples and the in vitro preimplantation single-cell objects were then processed and clustered

according to the same pipeline as for the STAMP atlas. Finally, preimplantation in vitro clusters were connected to in vivo clusters with labels predicted from running SingleR (v.1.8.1) on the in vivo cells with in vitro clusters as a reference. The predicted labels were then used to construct an alluvial plot using ggalluvial (v.0.12.3).

Statistical analysis and reproducibility

Statistical analysis was performed using two-tailed *t*-tests, Mann–Whitney *U*-tests or log-rank tests as indicated in the figure legends. $P < 0.05$ was considered to be statistically significant. All box and whisker plots demarcate the median (centre line), 75th and 25th percentiles (upper and lower bounds, respectively), and the minimum and maximum values excluding points determined to be outliers by exceeding 1.5× the interquartile range (whiskers). The bar graphs represent the mean ± s.d. or mean ± s.e.m. as indicated. Mice, tumour and cell numbers per condition are provided in the figure legends. Statistical analysis was performed using GraphPad Prism (v.9.4.1) or Python (v.3.10.3) using SciPy (v.1.8.0). Flow cytometry analysis was performed using FlowJo (v.10). The imaging experiments shown in Figs. 1a, 2a,f,g and 3d,h and Extended Data Figs. 1a–c, 2b,e,g, 3c,e,f, 6a,b,f,h,l,m and 9a,e depict representative images and the quantifications are the aggregate of biological replicates as indicated. Mouse experiments were reliably reproduced. Experiments were replicated independently at least twice unless otherwise stated in the legend. No statistical methods were used to determine sample size. Mice were randomized before treatment. Investigators were not blinded to allocation. Treatment experiments were performed blinded for automated high throughput analysis. NGS and Flow cytometry analysis on tumour biopsies were not performed blinded owing to the needs to track single tumour biopsies and pool them by similar features (genotype of the mouse, immune phenotypes, time). For all the experiments, analysis was objective.

Figure preparation

All data were assembled into figures with Adobe Illustrator 2022. Figs. 1a, 2c and 3g were created using BioRender (https://biorender.com/). R plots used native plotting capabilities of the aforementioned packages together

with ggplot2 (v.3.3.5), ggpubr (v.0.4.0) and ComplexHeatmap[42] (v.2.6.2) packages.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All mouse sequencing data are publicly available at the Gene Expression Omnibus (GSE222231) (whole-exome sequencing, bulk RNA-seq, scRNA-seq). The ICON7 and IMvigor210 human datasets have been previously published[11,13] and deposited. IMvigor210 RNA-seq data were deposited to the European Genome–Phenome Archive under accession number EGAS00001002556. ICON7 raw RNA-seq and clinical data were deposited to the European Genome–Phenome Archive under accession number EGAS00001003487. Source data are provided with this paper.

Code availability

The STAMP image segmentation and analysis pipeline was developed using custom scripts in FIJI and Python and is available at GitHub (https://github.com/kcarbone/STAMP).

References

1. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer–immune set point. *Nature* **541**, 321–330 (2017).

2. Galon, J. & Bruni, D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov.* **18**, 197–218 (2019).

3. Galon, J. et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960–

1964 (2006).

4. Bruni, D., Angell, H. K. & Galon, J. The immune contexture and immunoscore in cancer prognosis and therapeutic efficacy. *Nat. Rev. Cancer* **20**, 662–680 (2020).

5. Tumeh, P. C. et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).

6. Gibney, G. T., Weiner, L. M. & Atkins, M. B. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol.* **17**, e542–e551 (2016).

7. Herbst, R. S. et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014).

8. Gajewski, T. F. The next hurdle in cancer immunotherapy: overcoming the non-T-cell-inflamed tumor microenvironment. *Semin. Oncol.* **42**, 663–671 (2015).

9. Bachhav, Y. G., Heinrich, A. & Kalia, Y. N. Controlled intra- and transdermal protein delivery using a minimally invasive Erbium:YAG fractional laser ablation technology. *Eur. J. Pharm. Biopharm.* **84**, 355–364 (2013).

10. Chung, W.-J. et al. Kras mutant genetically engineered mouse models of human cancers are genomically heterogeneous. *Proc. Natl Acad. Sci. USA* **114**, E10947–E10955 (2017).

11. Mariathasan, S. et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).

12. Dominguez, C. X. et al. Single-cell RNA sequencing reveals stromal evolution into LRRC15$^+$ myofibroblasts as a determinant of patient response to cancer immunotherapy. *Cancer Discov.* **10**, 232–253 (2020).

13. Desbois, M. et al. Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer. *Nat. Commun.* **11**, 5583 (2020).

14. Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015).

15. Spranger, S. & Gajewski, T. F. Impact of oncogenic pathways on evasion of antitumour immune responses. *Nat. Rev. Cancer* **18**, 139–147 (2018).

16. Liu, X. & Peng, G. Mitochondria orchestrate T cell fate and function. *Nat. Immunol.* **22**, 276–278 (2021).

17. Buck, M. D. et al. Mitochondrial dynamics controls T cell fate through metabolic programming. *Cell* **166**, 63–76 (2016).

18. Li, W. & Zhang, L. Rewiring mitochondrial metabolism for $CD8^+$ T cell memory formation and effective cancer immunotherapy. *Front. Immunol.* **11**, 1834 (2020).

19. Araki, K. et al. Translation is actively regulated during the differentiation of $CD8^+$ effector T cells. *Nat. Immunol.* **18**, 1046–1057 (2017).

20. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

21. Buechler, M. B. et al. Cross-tissue organization of the fibroblast lineage. *Nature* **593**, 575–579 (2021).

22. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

23. Tavernari, D. et al. Nongenetic evolution drives lung adenocarcinoma spatial heterogeneity and progression. *Cancer Discov.* **11**, 1490–1507 (2021).

24. Angelova, M. et al. Evolution of metastases in space and time under immune selection. *Cell* **175**, 751–765 (2018).

25. Tauriello, D. V. F. et al. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* **554**, 538–543 (2018).

26. Tahtinen, S. et al. IL-1 and IL-1ra are key regulators of the inflammatory response to RNA vaccines. *Nat. Immunol.* **23**, 532–542 (2022).

27. Doerr, J. et al. Whole-brain 3D mapping of human neural transplant innervation. *Nat. Commun.* **8**, 14162 (2017).

28. Geraldo, S., Simon, A. & Vignjevic, D. M. Revealing the cytoskeletal organization of invasive cancer cells in 3D. *J. Vis. Exp.* https://doi.org/10.3791/50763 (2013).

29. Helou, S. M. et al. A relationship-based approach to improving clinical trials: the imCORE research network experience. *Contemp. Clin. Trials* **86**, 105861 (2019).

30. Powles, T. et al. Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. *Nat. Med.* **25**, 1706–1714 (2019).

31. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

32. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

33. Saunders, C. T. et al. Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012).

34. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).

35. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, gky955 (2018).

36. Fontes, M. & Soneson, C. The projection score—an evaluation criterion for variable subset selection in PCA visualization. *BMC Bioinform.* **12**, 307 (2011).

37. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

38. Liberzon, A. et al. The Molecular Signatures Database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

39. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

40. Henningsson, R., Moratorio, G., Bordería, A. V., Vignuzzi, M. & Fontes, M. DISSEQT—distribution-based modeling of SEQuence space time dynamics. *Virus Evol.* **5**, vez028 (2019).

41. Zappia, L. & Oshlack, A. Clustering trees: a visualisation for evaluating clusterings at multiple resolutions. *Gigascience* **7**, giy083 (2018).

42. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

Acknowledgements

We thank the members of the C.M., S.J.T. and I.M. laboratories for advice, discussions and reagents; B. Hough and R. Asuncion for mouse husbandry; R. Garcia-Gonzalez, J. Yamada and E. Chua for veterinary care; the members of the Genentech Flow cytometry core group for technical assistance; M. Nitschke for help with RNA extraction; B. Korin for help with rebuttal experiments; and the members of the Genentech postdoctoral program for support. This study was funded by Genentech/Roche.

Author information

Author notes

1. These authors contributed equally: Guadalupe Ortiz-Muñoz, Markus Brown, Catherine B. Carbone, Ximo Pechuan-Jorge

Authors and Affiliations

1. Genentech, South San Francisco, CA, USA

   Guadalupe Ortiz-Muñoz, Markus Brown, Catherine B. Carbone, Ximo Pechuan-Jorge, Vincent Rouilly, Henrik Lindberg, Alex T. Ritter, Qianbo Sun, Tess Nicotra, Shreya R. Mantri, Angela Yang, Jonas Doerr, Deepti Nagarkar, Spyros Darmanis, Benjamin Haley, Sanjeev Mariathasan, Yulei Wang, Thomas Wu, Lelia Delamarre, Zora Modrusan, Richard Price, Shannon J. Turley, Ira Mellman & Christine Moussion

2. Department of Computer Science, Stanford University, Stanford, CA, USA

   Gautham Raghupathi

3. IUCT, Institut Universitaire du Cancer de Toulouse, Toulouse, France

   Carlos Gomez-Roca

4. Department of Pathology, Clínica Universidad de Navarra, Pamplona, Spain

   Carlos Eduardo de Andrea

5. Sarah Cannon Research Institute, Nashville, TN, USA

   David Spigel

6. Department of Pharmacology, UCSD, San Diego, CA, USA

Johannes Schöneberg

7. Department of Chemistry & Biochemistry, UCSD, San Diego, CA, USA

Johannes Schöneberg

Contributions

G.O.-M., M.B., C.B.C. and X.P.-J. designed and performed the experiments, analysed and interpreted the data, and contributed to writing the manuscript. C.B.C., V.R., H.L., G.R., J.S., G.O.-M., M.B. and X.P.-J. developed the high-content image analysis pipeline to analyse tumour features over time. A.T.R., J.D., S.R.M., T.N. and A.Y. performed experiments. Z.M. and S.D. supervised NGS experiments. X.P.-J., T.W. and Q.S. analysed NGS experiments. L.D. gave input on the vaccination experiment. B.H. provided synthetic vectors. R.P. supervised the collection of clinical samples. Y.W. and S.M. shared clinical trial data. C.G.-R., C.E.A. and D.S. collected patient samples. S.J.T. shared mouse model and blocking reagents for in vivo experiments. S.J.T. and I.M. edited the manuscript. D.N. and I.M. allocated resources. C.M. conceived and supervised the study, interpreted the data and wrote the manuscript with input from all of the authors. All of the authors read and approved the final Article.

Corresponding author

Correspondence to Christine Moussion.

Ethics declarations

Competing interests

G.O.-M., M.B., C.B.C., X.P.-J., V.R., H.L., D.N., L.D., A.T.R., A.Y., J.D., Y.W., S.M., S.D., Z.M., T.W., Q.S., B.H., R.P., S.J.T., I.M. and C.M. are employees of Genentech/Roche.

Peer review

Peer review information

Nature thanks Ana-Maria Lennon-Duménil, Miriam Merad and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Implantation, inflammatory response, and growth analysis of STAMP microtumours.

a, Image of **(i)** ears after P.L.E.A.S.E.® Laser microporation, **(ii)** microporated ears covered with tumour cell suspension, **(iii)** microporated ear seeded with tumour cells and covered with Matrigel. **b**, Representative tumour arrays of different tumour cell lines, n ≧ 10 animals per group. **c**, Representative time course image series of orthotopic B16F10 model cell line growing in STAMP. n = 5 animals. **d**, Manual analysis of growth kinetics of individual B16F10 STAMP tumours using ellipsoid formula to calculate volumes over time (upper panel) and using tumour segmentation to calculate area over time (lower panel). n = 27 tumours, 2 animals pooled. **e**, Representative tumour segmentation performed on validation images. (i) Input images, (ii) manually generated classification mask, (iii) features extracted after the penultimate upsampling step, (iv) output segmentation mask. **f-h**, Validation of the high content image analysis pipeline. **f**, Manual analysis of growth kinetics of individual KPP-EGFP tumour volumes (mm^3) using the ellipsoid formula. **g**, Manual analysis of growth kinetics of individual tumour areas (mm^2) using tumour segmentation. **h**, Automated analysis of growth kinetics of individual tumour areas (mm^2), **f-h**, n = 72

tumours, 3 animals pooled. **i**, Lymphoid and myeloid immune cell profiling of STAMP tumour array and subcutaneous tumours of KPP-EGFP cancer cells 18 days post tumour implantation and pooled. n = 6 ears, n = 3 mice for STAMP. n = 5 mice for subcutaneous tumours. **j**, Lymphoid and myeloid immune cell profiling of PBS or KPP-EGFP-seeded micropores. STAMP tumour arrays were harvested 18 days post tumour implantation and pooled by ear. n = 16 ears, n = 8 mice. Data are mean +/− s.d. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. *P* values are shown in the graph. **k**, Principal Component Analysis (PCA) of the 200 most variable genes of the RNA-seq data from healthy skin, pores and KPP-EGFP microtumours biopsied at day 1, day 3 and day 8 after tumour seeding and/or laser poration. Samples are coloured according to day and experimental group (tumour, pores and healthy skin). Ellipses at 95% confidence level are shown for each group. The percentage of explained variance for each principal component is annotated on the principal component axes. **j**, Heatmap of the z-scored gene expression values for Tnf, Il6, Il1b. Columns are annotated as in **k** and rows are hierarchically clustered to show genes with similar patterns of expression across the samples.

Source Data

Extended Data Fig. 2 KPP STAMP microtumour rejection is mediated by tumour antigen specific CD8 T cells.

a, Automated analysis of growth kinetics of KPP-EGFP total tumour array area (mm^2) per animal in Rag-2-deficient and wild type animals as described in Fig. 1b, n = 5 animals per group. **b**, Representative images of KPP-EGFP tumour arrays in wild type (upper panel) or Rag-2-deficient (lower panel) mice at 6 and 14 days post tumour implantation related to **b**. Red encircled tumours are rejected and white encircled tumours are persistent between time points. **c**, Survival probability of individual tumours of KPP-EGFP arrays in mice treated with CD8 depleting (n = 318 tumours, 5 animals) or isotype control antibodies (n = 534 tumours, 5 animals). **d**, Automated analysis of growth kinetics of individual tumour area (mm^2) for experiment as described in Fig. 1d. Red lines indicate tumours that are rejected, grey lines indicate tumours that persist. n = 100 tumours, n = 5 animals in Rag-2-deficient animals reconstituted with WT T cells, n = 88 tumours, n = 5

animals in Rag-2-deficient animals reconstituted with OT-I T cells. **e**, Representative images of STAMP tumour arrays for experiment as described in **d**. Red = T cells, green = KPP-EGFP. **f**, T cell infiltration kinetics of individual tumours measured by tdTomato MFI of T cells for experiment as described in **d**. n = 5 animals per group, n = 61 tumours in Rag-2-deficient animals reconstituted with WT T cells and n = 53 tumours in Rag-2-deficient animals reconstituted with OT-I T cells. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **g**, Representative image time course of KPP-OVA-EGFP tumour arrays in Rag-2-deficient animals reconstituted with wild type or OT-I T cells. n = 4 animals reconstituted with WT T cells, n = 5 animals reconstituted with OT-I T cells, n = 156 tumours in animals reconstituted with WT T cells, n = 93 tumours in animals reconstituted with OT-I T cells. Red = T cells, green=KPP-EGFP. **h**, Survival probability of individual tumours from mice bearing KPP-EGFP-OVA tumour in Rag-2-deficient animals reconstituted with wild type or OT-I T cells and treated with CD8 depleting or isotype control antibody. n = 4 WT-OVA isotype, n = 5 OT-I-OVA isotype and n = 5 OT-I-OVA anti-CD8 treated animals per group, n = 93 OT-I-OVA isotype, n = 156 WT-OVA isotype and n = 206 OT-I-OVA anti-CD8 treated tumours. Statistical analysis was performed with log-rank test (referenced to WT-OVA isotype). **i**, Frequency of tumour antigen (M86) specific T cells in M86-encoding RNA-LPX vaccinated or naive T cell donor mice shown as PD-1+ percent of CD8+ activated memory T cells. **j**, Survival probability of individual tumours from mice bearing KPP-M86-mTagBFP2 microtumours in Rag-2-deficient mice reconstituted with tumour antigen specific T cells (M86) from vaccinated mice or T cells from naive mice. n = 73 tumours in vaccinated animals, n = 72 tumours in naive animals, 4 animals per group pooled. **c,f,h,i**, *P* values are shown in the graph and colour-coded for the treatment group in **h**. **c,h,j**, The centre line shows the Kaplan-Meier curve, the shaded area shows the 95% confidence interval. Statistical analysis was performed with log-rank test.

Source Data

Extended Data Fig. 3 Transcriptionally distinct lesions with inflamed, excluded, and desert immune phenotypes coexist in tumour arrays.

a, PCA of the top 1000 most variable genes of the bulk RNA-seq data of individually biopsied tumours with differential immune phenotype inflamed, desert, excluded. Ellipses at 95% confidence level are shown for each group. The percentage of explained variance for each principal component is annotated on the principal component axes. **b**, Detailed GSEA for selected pathways that are significantly enriched between three different immune phenotypes (inflamed, desert, excluded) from STAMP tumours. **c**, Representative STAMP tumour array of KPP-EGFP in Foxn1-Nude mice reconstituted with tdTomato + T cells at 8 days post tumour implantation. Magenta = T cells, cyan = KPP-EGFP. n = 5 animals. **d**, Proportion of immune phenotypes present in STAMP arrays of three murine tumour models. NSCLC = non-small cell lung cancer (n = 175 tumours, n = 2 animals), KPP = pancreatic ductal adenocarcinoma (n = 240 tumours, n = 4 animals), and B16F10 = melanoma (n = 229 tumours, n = 4 animals). **e**, Representative images of STAMP tumour arrays of KPP-EGFP on abdominal skin of Rag-2-deficient mice reconstituted with tdTomato T cells, 8 days post tumour implantation, Red = T cells, green = KPP-EGFP. Overview of the entire abdominal tumour array (left panel), enlarged images of neighbouring tumours with inflamed, excluded, and desert immune phenotypes (right panels). n = 3 animals. **f**, Representative images of experimental lung metastases of KPP-EGFP in Rag-2-deficient mice reconstituted with tdTomato T cells, 8 days post tumour implantation. Red = T cells, green = KPP-EGFP. Overview of the entire lung lobe (left panel), enlarged images of metastatic lesions with inflamed, excluded, and desert immune phenotypes (right panel). n = 3 animals.

Source Data

Extended Data Fig. 4 Inflamed, excluded, and desert immune phenotypes are not predetermined by genetic or non-genetic heterogeneity in the *in vitro* cultured cell line.

a, Histogram depicts variant allele frequency (VAF) detected by whole exome sequencing of the KPP-EGFP cell line before implantation. n = 3 replicates indicated by colour with a corresponding dashed line showing the median VAF of the distribution. Homozygous and the heterozygous peaks indicated by black dashed lines. **b**, Neutrality test of low frequency variants

for each replicate in **a**. The coefficient of determination (R^2) is shown to indicate goodness of fit. The ratio of the mutation rate (μ) and the effective division rate (β) is the slope of the least squares fitted line according to the neutral cumulative mutation distribution. **c**, UpsetR plot showing the number of potentially explanatory variants detected by whole exome sequencing of pooled tumour biopsies with different immune phenotypes (x-axis). Variants have been confirmed after manual curation using detailed inspection of the alignment reads with Integrative Genome Viewer. **d**, UMAP embedding showing the KPP tumour cell clusters of scRNA sequencing data from *in vitro* cultured cells pre-implantation in STAMP. **e**, UMAP embedding showing the KPP tumour cell clusters of scRNA sequencing data from *in vivo* STAMP tumours post-implantation. **f**, Alluvial plot depicting the relationship between single cells from KPP-EGFP post implantation and its corresponding phenotypes together with the most likely *in vitro* cluster assignment as determined by SingleR.

Extended Data Fig. 5 Desert, excluded or inflamed phenotypes are infiltrated by similar proportions of T cell subsets and clonotypes.

a, UMAP embedding of tdTomato+ T cell subclusters (indicated by colours) from pooled STAMP tumour biopsies. **b**, Dot plot showing the relative expression of important marker genes within T cell subclusters. Relative expression level indicated by colour, and percent of cells expressing the transcript indicated by circle size. **c**, Relative abundance for each T cell subcluster separated by immune phenotype. **d**, Flow cytometry-based T-lymphoid immune cell profiling of rejected, inflamed, excluded and desert tumours at day 10. Absolute frequency of CD3, CD4 and CD8 T cells (upper plot) and proportion of naive, activated/resident, effector memory, and central memory of CD4 or CD8 T cells (lower plot). $n = 6$ animals with 3 tumours pooled per phenotype per animal. Data are mean +/−s.d. Statistical analysis was performed using a two-tailed t-test. *P* values are shown at the top of the graph. **e**, TCR clonotype diversity indices for each immune phenotype. **f**, Cell number of each T cell clonotype with cluster identification. **g**, Pathway analysis comparing excluded vs inflamed for combined top clonotypes. p-values were fdr adjusted and reported if p-value adjusted < 0.1.

Source Data

Extended Data Fig. 6 Spatial T cell patterns dictate efficiency of anti-tumour attack.

a-b, Representative *in vitro* time-lapse image series of KPP-GCaMP6-Her2-expressing tumour organoids cultured with TdTomato T cells in the absence **(a)** or presence **(b)** of anti-Her2/anti-CD3 T cell-dependent bispecific antibody (TDB). Green = GCaMP6, red = propidium iodide, magenta = T cells. n = 3 biological replicates. Representative trace showing flashing behaviour (delta MFI) of GCaMP6 (blue) and influx of PI (red) fluorescence over time in the absence (control, light colour) or presence (TDB, dark colour) of TDB as described in **a** and **b. c**, Average GCaMP6 deltaMFI (blue) and average PI deltaMFI (red) plotted against the time (min) in the absence (control, light colour) or presence (TDB, dark colour) of TDB as described in **a** and **b**. Data are mean +/− s.e.m. **d**, GCaMP6 flashing index in the absence (control) or presence of TDB as described in **a** and **b**. Data are mean +/− s.e.m. Statistical analysis was performed using a one-tailed Mann-Whitney U-test. **e**, PI influx in the absence (control) or presence of TDB as described in **a** and **b**. Data are mean +/− s.e.m. Statistical analysis was performed using a one-tailed Mann-Whitney U-test. **f**, Representative *in vivo* 2-photon time-lapse images of KPP-GCaMP6-mTagBFP2-Her2 STAMP tumours in Rag-2-deficient mice 12 days after tumour cell seeding and adoptive transfer of tdTomato+ T cells. Imaging before intravenous (I.V.) administration of TDB (upper panel, n = 3 animals), and after I.V. administration of TDB (lower panel, n = 2 animals). Green = GCaMP6 (KPP), blue = mTagBFP2 (KPP), red = tdTomato (T cells). **g**. GCaMP6 flashing index of STAMP tumours before TDB or after TDB administration, as described in **f**. Green = GCaMP6 (KPP), blue = mTagBFP2 (KPP), red = tdTomato (T cells). The centre line shows the median, the box limits show the minimum and maximum values. **h**, Representative in vivo 2-photon time-lapse images of KPP-GCaMP6-mtagBFP2-Her2 STAMP tumours in Rag-2-deficient mice with (upper panel) or without (lower panel) reconstitution with tdTomato + T cells and 12 days after tumour cell seeding. Green = GCaMP6 (KPP), blue = mTagBFP2 (KPP), red = tdTomato (T cells). n ≧ 3 animals. **i**, GCaMP6 flashing index of STAMP tumours in Rag-2-deficient mice with and without adoptive transfer of tdTomato+ T

cells, as described in **h**. The centre line shows the median, the box limits show the minimum and maximum values. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **j**, T cell abundance (MFI) of inflamed and excluded KPP-GCaMP6-mtagBFP2-Her2 STAMP tumours in Rag-2-deficient mice 8 days after tumour cell seeding and adoptive transfer of tdTomato+ T cells. n = 6 excluded and n = 15 inflamed tumours. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **k**, GCaMP6 flashing index of inflamed and excluded STAMP tumours, as described in **j**. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **l**, Representative in vivo epifluorescence time-lapse images of GCaMP6-expression of inflamed KPP tumours. n = 15 tumours. Green = GCaMP6. **m**, Representative *in vivo* epifluorescence time-lapse images of GCaMP6-expression of excluded KPP tumours. n = 6 tumours. Green = GCaMP6. **d, e, i, j, k**, P values are shown at the top of the graph.

Source Data

Extended Data Fig. 7 Myeloid depletion reduces T cell recruitment and transition towards inflamed phenotype.

a, STAMP scRNA-seq atlas UMAP embedding of innate, myeloid, and stromal cell clusters from STAMP single tumour biopsies pooled per immune phenotype. **b**, UMAP embedding for the myeloid compartment of the scRNA from single tumour biopsies pooled per immune phenotypes. **c**, Dot plot for expression of Cxcl9 and Cxcl10 across all cell clusters of the scRNA sequencing STAMP tumour atlas described in **a**. **d**, Flow cytometry-based myeloid cell profiling of 5 pooled tumour biopsies per immune phenotype at 10 days post tumour implantation. n = 6 animals. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. *P* values are shown at the top of the graphs. **e**, Scores of myeloid signatures in bulk RNA-seq of tumours at 8 days post-implantation. n = 7 tumours per group (gene signatures for each myeloid population have been derived from the scRNA-seq myeloid STAMP Atlas). Box plots show quartiles of the dataset and whiskers show the rest of the distribution. Statistical analysis was performed using a two-tailed t-test with unequal variances, adjusted for false discovery rate. P values are at the top of the

graphs. **f**. Total T cell abundance in individual tumours over time in control, Ly6C- and Gr1-depleted animals as described in Fig. 3b, n = 6 Isotype control-treated animals, n = 5 Ly6c- and n = 5 Gr1-depleted animals, n = 319 Isotype control-treated, n = 301 Ly6c- and n = 187 Gr1-depleted tumours. Box plots show quartiles of the dataset and whiskers show the rest of the distribution, excluding points determined to be outliers by exceeding 1.5x the interquartile range. Statistical analysis was performed using a two-tailed t-test (referenced to Isotype control treated tumours). P values are shown in the graph. **g**, Analysis of immune phenotype transition dynamics of tumours in Isotype control-treated, Ly6C- and Gr1-depleted animals related to **f**. Arrows represent fold changes in transition likelihood between control and depleted animals. Black = increased transition, blue = decreased transition.

Source Data

Extended Data Fig. 8 Depletion of Dpt+ fibroblasts promotes immune desert phenotype in STAMP.

a, UMAP embedding for the fibroblast compartment for the scRNA from pooled single tumour STAMP biopsies. **b**, Dermatopontin (Dpt) expression overlaid on the fibroblast UMAP. **c**, Violinplot showing the expression levels of Dpt across fibroblast subclusters. **d**, Barplot showing the relative abundance of the fibroblast subclusters for microporated skin. **e**, Quantification of fibroblast frequency in skin of Dpt-CreERT2_LSL-DTR-YFP mice with and without Tamoxifen and/or Diphtheria toxin (DTX) treatment. n = 4 Tamoxifen/DTX-treated Control mice, n = 4 Tamoxifen-treated DTR-YFP mice and n = 5 Tamoxifen/DTX-treated DTR-YFP mice. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **f**, Total T cell abundance in individual tumours over time in control and Dpt-depleted animals as described in Fig. 3h. n = 5 control and n = 7 Dpt-depleted animals, n = 207 control and n = 314 Dpt-depleted tumours. Box plots show quartiles of the dataset and whiskers show the rest of the distribution, excluding points determined to be outliers by exceeding 1.5x the interquartile range. Statistical analysis was performed using a two-tailed t-test. **g**, Analysis of immune phenotype transition dynamics of tumours in Dpt-depleted animals related to **f**. Arrows represent fold changes in transition likelihood between control and depleted animals.

Black = increased transition, Blue = decreased transition. **h**, Dot plot for expression of ligand-receptor pairs of the main secreted factors of the ChemoCAF for all of the fibroblast and myeloid populations from the scRNA sequencing STAMP tumour atlas. **i-j**, Communication probability estimated by CellChat for **i**, CXCL chemokine pathway with fibroblasts subclusters as the sender populations and myeloid subclusters as the receiver and **j**, CCL chemokine pathway with fibroblasts subclusters as the sender populations and myeloid subclusters as the receiver. **k-l**, Flow cytometry-based immune cell profiling of tumours in Dpt-depleted versus non-depleted animals. n = 6 control animals and n = 8 DTR animals. Data are mean +/− s.e.m. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. **k**, Shows cell types with no significant change upon depletion. **l**, Shows cell types with significant differences between control and depletion. **e,f,k,l**, *P* values are shown at the top of the graphs.

Source Data

Extended Data Fig. 9 Early transition to an inflamed phenotype predicts tumour response to immunotherapy.

a, Histology of patient paired tumour biopsies at baseline and at progression after treatment with checkpoint blockade. CD8 staining in brown, pan-cytokeratin staining in magenta. **b**, Individual tumour growth kinetics of KPP-GFP STAMP tumours for experiment with immune reconstituted animals described in Fig. 4a. **c-d**, STAMP arrays of KPP-EGFP were implanted in immunocompetent animals and treated 10 days post-implantation with isotype control antibodies (n = 130 tumours, 4 animals), anti–PD-L1 (n = 94 tumours, 4 animals), anti–TGF-β (n = 61 tumours, 3 animals) or a combination of anti–PD-L1 with anti–TGF-β (n = 99 tumours, 4 animals). **c**, Individual tumour growth kinetics (mm^2) shown with coloured lines to indicate tumours that are rejected and grey lines to indicate tumours that persist. **d**, Survival probability of individual tumours described for **c** and **d**. The centre line shows the Kaplan-Meier curve, the shaded area shows the 95% confidence interval. Statistical analysis was performed with log-rank test. *P* value is shown in the graph and colour-coded for the treatment group. **e**, Image series of individual STAMP tumour over time. Red = T cells, blue = tumour cells. **f**, Quantification of median radial fluorescence profile for

individual tumour shown in **e** to distinguish non-desert, T cell excluded phenotype (low T cell infiltration ratio, day 11) and T cell inflamed phenotype (high T cell infiltration ratio, day 15). **g**, Overlay of the automated classification of immune phenotypes with the individual tumour growth curve for the example in **e**. **h**, Hierarchical clustering of individual tumour immune growth rates (left) and T cell infiltration (right) for isotype control treated tumours and combo anti–PD-L1/anti–TGF-β treated tumours as described in Fig. 4a. Black = tumour resolved, white = mouse death/euthanasia. Cyan=combination treated responders (C.R. + P.R.), magenta = combination treated non-responders (S.D. + P.D.), blue = Control responders (C.R.+ P.R.), red = control non-responders (S.D. + P.D.).

Source Data

Extended Data Fig. 10 Early transition to an inflamed phenotype predicts tumour response to immunotherapy.

a-c, Total T cell abundance (tdTomato MFI) for individual microtumours as described in Fig. 4a. Isotype control antibodies (n = 554 tumours, 9 animals, grey dots), **a**, anti–PD-L1 (n = 399 tumours, 6 animals, yellow dots), **b**, anti–TGF-β (n = 287 tumours, 5 animals, orange dots), or a **c**, combination of anti–PD-L1 with anti–TGF-β (n = 642 tumours, 11 animals, red dots). Box plots show quartiles of the dataset and whiskers show the rest of the distribution, excluding points determined to be outliers by exceeding 1.5x the interquartile range. Statistical analysis was performed using a two-tailed Mann-Whitney U-test. P values are shown in the graphs. **d**, Immune-histories of individual KPP-GFP STAMP tumours (grey lines) implanted in E8I.CD8A.Cre-Rosa26.LSL.tdTomato animals and treated at day 10 post-implantation with isotype control antibodies or combo anti-PD-L1/anti-TGF-b. n n ≧ 86 tumours per group and 3-4 animals per group. Average tumour trajectory changing T cell abundance (y-axis) and inflammation ratio (x-axis) over time for each treatment condition is shown in bold, with colour scale indicating time. **e**, Immune-histories for individual tumours of combination treated animals grouped as responders or non-responders for experiment described in Fig. 4a. Average immune-history for each class is shown in bold, with colour scale indicating time. **f**, Graphical summary of immune dynamics leading to tumour rejection or progression.

Source Data

Supplementary information

Supplementary Information

Supplementary Figs. 1–5.

Reporting Summary

Supplementary Table 1

Differentially expressed genes ($q \leq 0.05$) between individual STAMP tumours of different immune phenotypes (inflamed, desert, excluded) (related to Extended Data Fig. 3a,b).

Supplementary Table 2

Normalized enrichment scores for all pathways that show significance in at least one comparison between the ICON 7 and IMvigor210 human clinical trials and STAMP mouse tumours—inflamed, excluded and desert STAMP tumours share common gene expression programs with clinical tumours of analogous immune phenotype (related to Fig. 2b).

Supplementary Table 3

Top 10 marker genes defining T cell subsets in STAMP microtumours (related to Fig. 2c,d and Extended Data Fig. 5a–c,f).

Supplementary Table 4

Differentially expressed genes by T cell subsets in inflamed versus excluded STAMP microtumours (related to Extended Data Fig. 5g).

Supplementary Table 5

Top10 marker genes defining myeloid cell subsets in STAMP microtumours (related to Fig.3a and Extended Data Fig. 7b,c).

Supplementary Table 6

Top 10 marker genes defining fibroblast subsets in STAMP microtumours (related to Fig.3f and Extended Data Fig. 8a–d).

Supplementary Table 7

Clinical summary (related to Extended Data Fig. 9a).

Supplementary Table 8

Markov probabilities of the transition between immune phenotypes (related to Fig. 4d).

Supplementary Video 1

Example of live in vivo two-photon imaging of an individual KPP BFP^+ STAMP tumour showing the migration of $TdTomato^+$ OTI T cells and WT GFP^+ T cells.

Supplementary Video 2

Confocal imaging followed by 3D reconstruction of a whole cleared lung showing KPP-GFP^+ tumour foci and $TdTomato^+CD3^+$ T cells 8 days after i.v. transfer (related to Extended Data Fig. 3f).

Supplementary Video 3

Confocal live imaging showing low Ca^{2+} flashes (green) in a $GCaMP6^+HER2^+$ KPP organoid, co-cultured with T cells (magenta) and propidium iodide (Red) (related to Extended Data Fig. 6a).

Supplementary Video 4

Confocal live imaging showing high Ca^{2+} flashes (green) in a $GCaMP6^{+}HER2^{+}$ KPP organoid, co-cultured with T cells (magenta) and TDB (anti-HER2/anti-CD3) to force tumour cell killing by T cells (positive control). Propidium iodide entry (red) reveals tumour cell death (related to Extended Data Fig. 6b).

Supplementary Video 5

In vivo two-photon live imaging showing absence of Ca^{2+} flashes in $GCaMP6^{+}BFP^{+}HER2^{+}$ KPP tumours implanted in nude mice (Related to Extended Data Fig. 6h).

Supplementary Video 6

In vivo two-photon live imaging showing some Ca^{2+} flashes in $GCaMP6^{+}BFP^{+}HER2^{+}$ KPP tumours implanted into nude mouse reconstituted with $TdTomato^{+}$ T cells (related to Extended Data Fig. 6f,h).

Supplementary Video 7

In vivo two-photon live imaging showing high level of Ca^{2+} flashes in $GCaMP6^{+}BFP^{+}HER2^{+}$KPP tumours implanted in nude mice reconstituted with $TdTomato^{+}$ T cells and treated with TDB (anti-HER2/anti-CD3) to force tumour cell killing by T cells (related to Extended Data Fig. 6f).

Supplementary Video 8

Epifluorescence live imaging of an immune inflamed tumour showing high Ca^{2+} flashes in $GCaMP6^{+}$ KPP tumour cells implanted in nude mice reconstituted with $TdTomato^{+}$ T cells (related to Extended Data Fig. 3l).

Supplementary Video 9

Epifluorescence live imaging of an immune-excluded tumour showing low Ca^{2+} flashes (green) in $GCaMP6^{+}$ KPP tumour cells implanted in nude mice reconstituted with $TdTomato^{+}$ T cells (related to Extended Data Fig. 3m).

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 9

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to

the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Ortiz-Muñoz, G., Brown, M., Carbone, C.B. *et al.* In situ tumour arrays reveal early environmental control of cancer immunity. *Nature* **618**, 827–833 (2023). https://doi.org/10.1038/s41586-023-06132-2

- Received: 03 May 2021
- Accepted: 25 April 2023
- Published: 31 May 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06132-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06132-2

| Section menu | Main menu |

- Article
- Open Access
- Published: 07 June 2023

Ultraviolet radiation shapes dendritic cell leukaemia transformation in the skin

- Gabriel K. Griffin ,
- Christopher A. G. Booth,
- Katsuhiro Togami,
- Sun Sook Chung,
- Daniel Ssozi,
- Julia A. Verga,
- Juliette M. Bouyssou,
- Yoke Seng Lee,
- Vignesh Shanmugam,
- Jason L. Hornick,
- Nicole R. LeBoeuf,
- Elizabeth A. Morgan,
- Bradley E. Bernstein,
- Volker Hovestadt ,
- Peter van Galen &
- …
- Andrew A. Lane

Show authors

Nature **volume 618**, pages 834–841 (2023)

- 8036 Accesses

- 1 Citations
- 263 Altmetric
- Metrics details

Abstract

Tumours most often arise from progression of precursor clones within a single anatomical niche. In the bone marrow, clonal progenitors can undergo malignant transformation to acute leukaemia, or differentiate into immune cells that contribute to disease pathology in peripheral tissues[1,2,3,4]. Outside the marrow, these clones are potentially exposed to a variety of tissue-specific mutational processes, although the consequences of this are unclear. Here we investigate the development of blastic plasmacytoid dendritic cell neoplasm (BPDCN)—an unusual form of acute leukaemia that often presents with malignant cells isolated to the skin[5]. Using tumour phylogenomics and single-cell transcriptomics with genotyping, we find that BPDCN arises from clonal (premalignant) haematopoietic precursors in the bone marrow. We observe that BPDCN skin tumours first develop at sun-exposed anatomical sites and are distinguished by clonally expanded mutations induced by ultraviolet (UV) radiation. A reconstruction of tumour phylogenies reveals that UV damage can precede the acquisition of alterations associated with malignant transformation, implicating sun exposure of plasmacytoid dendritic cells or committed precursors during BPDCN pathogenesis. Functionally, we find that loss-of-function mutations in *Tet2*, the most common premalignant alteration in BPDCN, confer resistance to UV-induced cell death in plasmacytoid, but not conventional, dendritic cells, suggesting a context-dependent tumour-suppressive role for TET2. These findings demonstrate how tissue-specific environmental exposures at distant anatomical sites can shape the evolution of premalignant clones to disseminated cancer.

Main

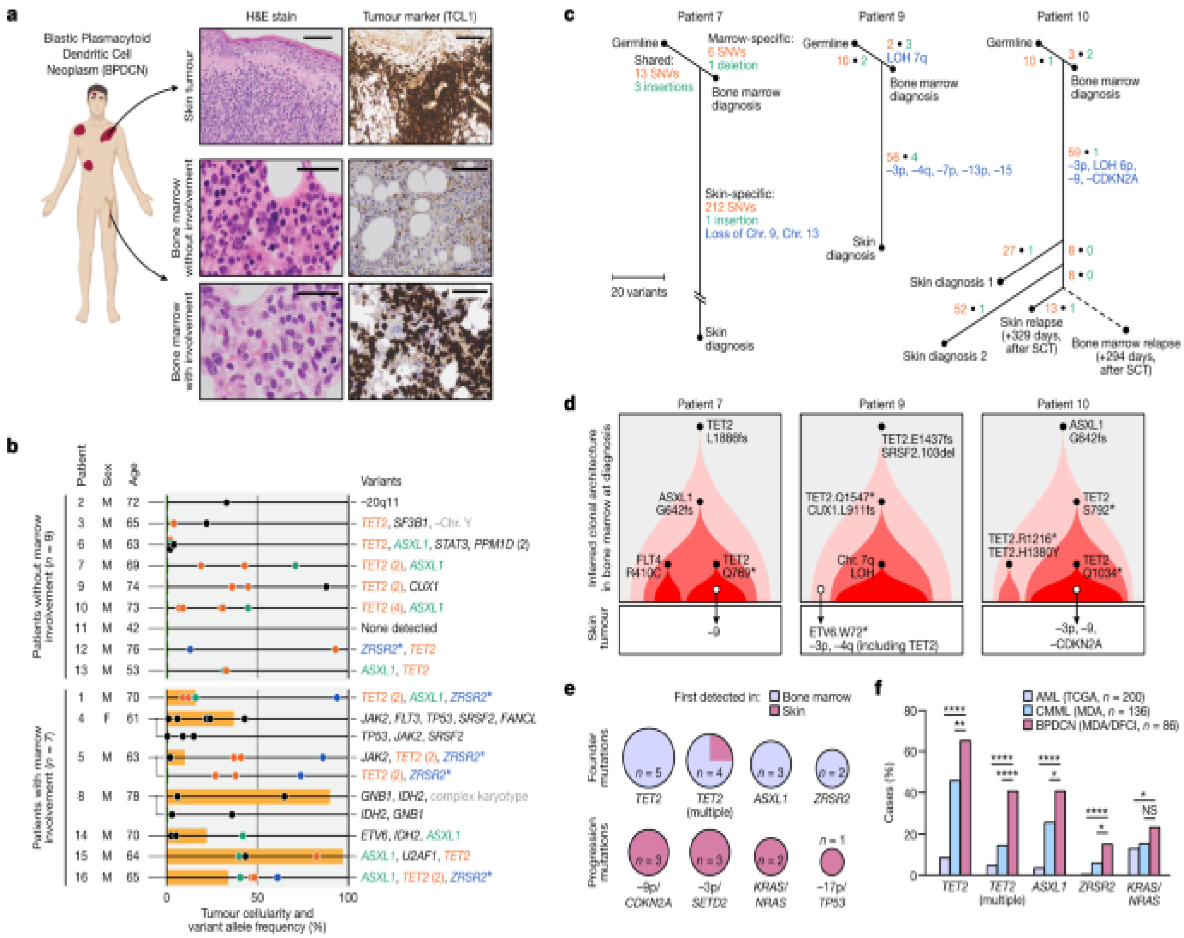
Clonal expansions of cells containing somatic mutations are common in normal tissues, and often arise during ageing or in response to genotoxic

stress[6,7,8]. Although most clones never progress, rare cells acquire additional alterations that confer a proliferative or survival advantage within the local tissue environment. In the bone marrow of ageing individuals, clonally expanded precursors with preleukaemic mutations give rise to a variety of differentiated immune cell populations that circulate throughout the body in the peripheral blood[1,2,3]. These cells can migrate into peripheral tissues and are increasingly recognized as mediators of organ-specific inflammation[9,10,11]. Whether tissue-specific mutational processes affect the clonal evolution of these immune cells is poorly understood. Here we used integrated genomic and single-cell analysis to examine progression to malignancy in BPDCN, an aggressive leukaemia that often presents as isolated skin tumours without clinically apparent blood or marrow involvement[12]. Thus, BPDCN presents a unique opportunity to study clonal evolution to cancer across anatomical sites, and to evaluate tissue-specific mutational processes in malignant transformation and progression.

Phylogenomics of BPDCN across tissues

BPDCN is a unique acute leukaemia that often presents with malignant cells isolated to the skin[13,14] (Fig. 1a). Despite this localized presentation, most patients ultimately develop systemic disease[12,15]. Although these features suggest that BPDCN may arise in the skin, other findings favour a bone marrow origin. For example, BPDCN is frequently associated with underlying clonal haematopoiesis and can also present with concurrent bone marrow and blood involvement[12,14,15,16]. These observations raise fundamental questions about the anatomical origins, ontogeny and pathogenesis of BPDCN.

Fig. 1: Phylogenomics of BPDCN across tissues.



a, Skin tumour (top) from a representative patient with BPDCN showing infiltration by malignant cells (haematoxylin and eosin (H&E), left) expressing the pDC marker TCL1 (immunohistochemistry, right). Marrow can show normal haematopoiesis (middle) or involvement by malignant cells (bottom). Scale bars, 200 μm (top), 100 μm (middle right and bottom right), 50 μm (middle left and bottom left). **b**, Mutated genes identified by targeted sequencing of marrow samples from 16 patients with BPDCN, including nine without marrow involvement (top; limit of detection, 5%) and seven with involvement (bottom; the orange bars indicate tumour cellularity). Gene labels are ordered from left to right from low to high VAF (*x* axis). The parentheses indicate multiple mutations in one gene. Recurrent gene mutations (≥3 patients) are indicated in colour. The asterisks indicate genes on chromosome (Chr.) X. **c**, Tumour phylogenies for patients 7, 9 and 10. Single-nucleotide variants (SNVs) (red), insertions/deletions (green) and copy-number alterations (blue, minus symbol represents chromosome or arm losses) are indicated (Methods). The dashed line indicates a post-transplant

sample for which new alterations could not be detected owing to donor SNVs. SCT, stem cell transplant. **d**, Inferred clonal architecture in marrow samples for patients in **c**. Subclones directly related to BPDCN skin tumours are indicated by arrows, and are further annotated with skin-specific progression mutations. **e**, The frequency of mutations first detected in marrow (founder mutations, top) or skin tumours (progression mutations, bottom) for patients in **c** and Extended Data Fig. 1d,e. *n* values indicate the number of gene mutations across the five patients. **f**, The frequency of mutations in founder and progression genes in BPDCN, CMML and AML. Statistical significance was determined using two-sided Fisher's exact tests. DFCI, Dana-Farber Cancer Institute; MDA, MD Anderson Cancer Center; TCGA, The Cancer Genome Atlas. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{****}P < 0.0001$; NS, not significant. The diagram in **a** was created using BioRender.

Source Data

To define the relationship between clonal (premalignant) bone marrow and BPDCN skin tumours, we identified 16 patients with available biopsy material for sequencing (Supplementary Table 1). This cohort comprised nine patients without bone marrow involvement and seven patients with concurrent skin and marrow involvement at diagnosis, including three who achieved remission after initial treatment (Fig. 1b). Skin tumours showed hallmark features of BPDCN, including purple nodules, plaques or bruise-like patches, and expression of canonical diagnostic markers (CD123, TCL1, CD4 and CD56; Extended Data Fig. 1a–c). Patients were predominantly male (15 out of 16, 94%) and over the age of 60 (14 out of 16, 88%), consistent with the known male bias and age association of BPDCN[13,14].

Targeted sequencing of 23 bone marrow samples from this cohort revealed pathogenic mutations in 15 out of 16 (94%) patients (Fig. 1b and Supplementary Table 2a). The mutations affected genes that are recurrently altered in clonal haematopoiesis, myeloid leukaemia and BPDCN, including *TET2* (11 out of 16 patients, 69%), *ASXL1* (9 out of 16 patients, 56%) and RNA splicing factors (9 out of 16 patients, 56%)[1,2,3,12,17,18]. Mutations in these genes were identified at a high variant allele fraction (VAF) in 8 out of 9 (89%) patients without marrow involvement, and in 3 out of 3 (100%) patients at the time of marrow remission, suggesting their presence in clonally expanded progenitors. There was no significant difference in VAF

for samples with and without marrow involvement (34.9% and 29.2% for *TET2*, respectively, $P = 0.635$ by two-sided Student's *t*-test; 33.9% and 28.7% for *ASXL1*, respectively, $P = 0.672$).

We next defined the relationship between expanded premalignant haematopoietic clones in the bone marrow and BPDCN tumours in the skin at diagnosis using whole-exome (WES; Fig. 1c and Supplementary Table 2b–d), whole-genome (WGS; Extended Data Fig. 1d and Supplementary Table 2e) and targeted (Extended Data Fig. 1e) sequencing. This revealed a direct relationship between premalignant haematopoietic clones in the bone marrow and BPDCN skin tumours in 9 out of 9 cases (100%) with matched sequencing. Reconstruction of tumour phylogenies showed a modest number of clonally expanded somatic mutations in bone marrow (WES range, 16–23; WGS range, 161–416) and a higher number in BPDCN skin tumours (WES range, 74–229; WGS range, 759–2,798). Most bone marrow mutations were found in matched skin tumours (WES range, 68.8–70.6%), while the majority of BPDCN skin tumour mutations were unique to the skin (range, 83.8–93.1%; Fig. 1c). Likewise, copy-number alterations, such as loss of *CDKN2A*, were detected in skin tumours but not matched bone marrow samples (Fig. 1c and Extended Data Figs. 1d,f, 2, 3a and 4a–c). These findings identify bone marrow as the source of premalignant clones in BPDCN, and indicate mutational evolution at multiple tissue sites during malignant transformation.

To examine clonal evolution over time, we extended our tumour phylogenies by sequencing six relapse samples from three patients (patients 1, 10 and 12; Fig. 1c and Extended Data Figs. 1d, 3b and 4d,e). Each of these patients experienced an initial response to systemic therapy that was followed (months later) by relapse in the skin and involvement of bone marrow and blood. Most mutations present in skin tumours from patient 10 at diagnosis were also detected in relapse skin tumour and bone marrow samples collected 1 year later (59.8–71.7%; Fig. 1c). Moreover, relapse samples accumulated additional alterations that were not detected at diagnosis. For example, out of 6,071 mutations detected in relapse samples from patient 1 (collected 2 years after diagnosis), 46.1% were shared with the initial skin tumour, while 30.4% were specific to the latest relapse sample (Extended Data Fig. 1d). This indicates that malignant cells present at diagnosis persist through initial therapy to initiate disease relapse.

Clonal architecture and drivers of BPDCN

We next sought to define the clonal architecture and molecular drivers of BPDCN from precursor cells in the marrow to malignant tumours in the skin. We used VAFs to define founder clones and progression mutations in five patients (patients 1, 7, 9, 10 and 12) profiled by WES or WGS.

In contrast to clonal haematopoiesis of indeterminate potential[1,2,3], we observed complex clonal hierarchies in marrow samples from these patients. For example, marrow from patient 10 showed a truncal mutation in *ASXL1* and two subclones defined by distinct pairs of *TET2* mutations (subclone 1: S792* and Q1034*, subclone 2: R1216* and H1380Y; asterisk indicates gain of stop codon; Fig. 1d). Only mutations from subclone 1 were detected in skin tumours from patient 10, while subclone 2 remained isolated to the marrow. Likewise, marrow from patient 9 contained a single dominant clone with truncal *TET2* (E1437fs; frameshift) and *SRSF2* (P95–R102del; deletion) mutations, and subclonal events in *TET2* (Q1547fs), *CUX1* (L911fs) and 7q (CN-LOH; copy-number neutral loss of heterozygosity; Fig. 1d). Only the truncal *TET2* and *SRSF2* mutations were identified in the matched skin tumour, suggesting derivation from an ancestral clone that diverged before alterations that remained isolated to the marrow. Similar patterns of clonal evolution were observed in patients 1, 7 and 12 (Fig. 1d and Extended Data Fig. 1g). Overall, founder clones contained recurrent alterations in *TET2* (5 out of 5 cases, 4 out of 5 cases with ≥2 alterations), *ASXL1* (3 out of 5 cases) and splicing-factors (3 out of 5 cases), including *ZRSR2* (2 out of 5 cases), suggesting privileged roles in premalignant evolution in BPDCN (Fig. 1e (top)).

We next examined progression events associated with transformation of founder clones in the marrow to BPDCN tumours in the skin. Activating RAS mutations were detected in skin tumours in 2 out of 5 patients (patients 1 and 12) but not founder clones in the marrow (Fig. 1e (bottom)). Likewise, copy-number loss of tumour suppressor regions was detected in skin tumours but not the marrow, including *CDKN2A*/chromosome 9p (loss of 9p harbouring the tumor suppressor *CDKN2A*) in 3 out of 5 patients (patients 1, 7 and 10), *SETD2*/chromosome 3p in 3 out of 5 patients (patients 1, 9 and 10) and *TP53*/chromosome 17p in 1 out of 5 patients (patient 1; Fig. 1e

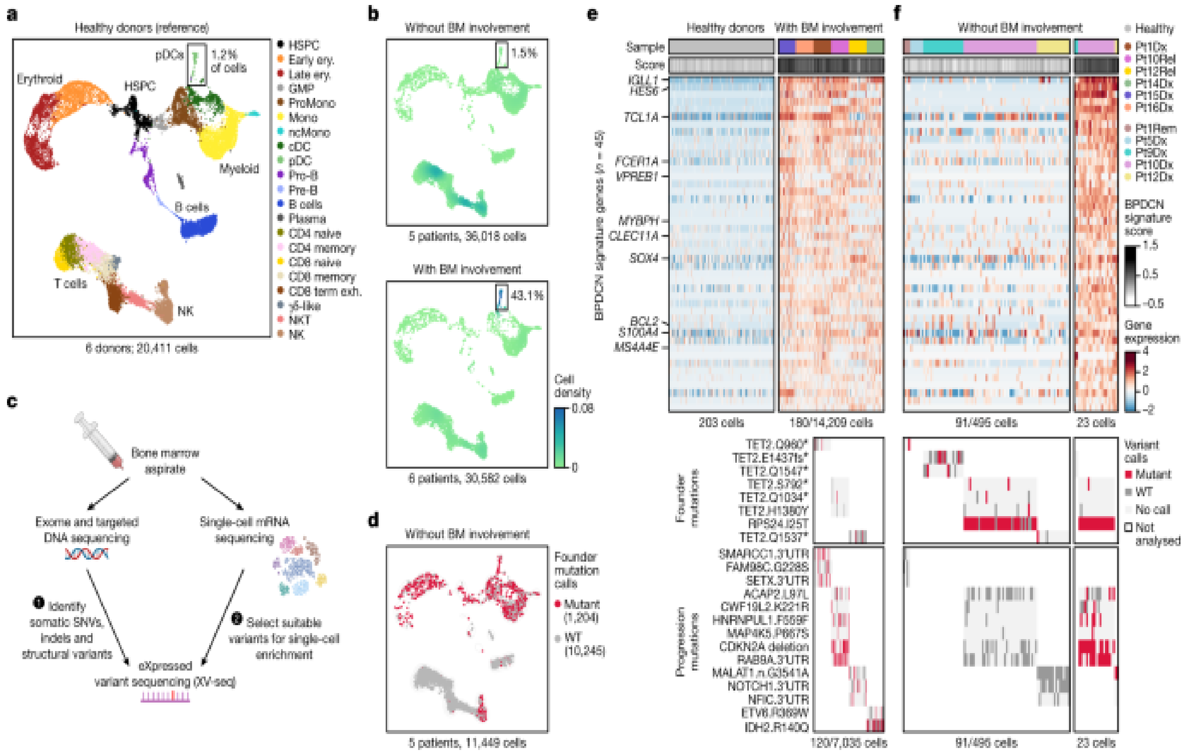
(bottom)). This supports key roles for these alterations in BPDCN transformation.

Finally, we compared the frequency of founder and progression-type mutations in two published BPDCN cohorts[12,17] to other myeloid neoplasms with similar clinicopathologic features, including chronic myelomonocytic leukaemia (CMML) and acute myeloid leukaemia (AML)[19,20,21,22]. Mutations in founder (*TET2*, *ASXL1* and *ZRSR2*) and progression (*NRAS*; *KRAS*; loss of chromosome 9 or 3p harbouring *CDKN2A* or *SETD2*, respectively) genes were all more frequent in BPDCN than CMML and/or AML, supporting privileged roles in BPDCN pathogenesis (Fig. 1f and Extended Data Fig. 1h). Notably, *TET2* alterations are present in nearly 70% of BPDCN tumours, including probable biallelic mutations in the majority of cases[12,17]. Moreover, approximately 10% of cases of BPDCN have *IDH2* hotspot mutations, which typically occur in cases with wild-type *TET2* and cause *TET2* inhibition through the 2-HG oncometabolite[12,17,23,24,25]. This supports functional inactivation of *TET2* in 70–80% of cases of BPDCN—to our knowledge, among the highest frequency of any cancer. This supports a unique role for *TET2* inactivation during BPDCN development.

Founder mutations across haematopoiesis

Our phylogenomic analysis indicated that BPDCN originates from clonally expanded progenitors in the bone marrow, consistent with recent findings[12,15,16]. To test this directly, we applied integrated single-cell RNA-sequencing (scRNA-seq) and genotyping analysis. We first generated a reference dataset from marrows of six healthy donors. This yielded 20,411 high-quality transcriptomes and represented the full spectrum of haematopoietic progenitor cells to differentiated myeloid, erythroid and lymphoid lineages (Fig. 2a and Supplementary Table 3). We next profiled marrow samples from five patients with BPDCN without marrow involvement (disease only in the skin) and six patients with involvement (disease in skin and the marrow; Fig. 2b). This yielded 66,600 high-quality transcriptomes.

Fig. 2: Single-cell profiling resolves premalignant pDCs and BPDCN.



a, Uniform manifold approximation and projection (UMAP) representation of the scRNA-seq analysis of marrow ($n = 20{,}411$ cells) from six healthy donors. Clusters show expected cell types, including progenitor, myeloid, erythroid and lymphocyte-lineage cells and pDCs (box). HSPC, hematopoietic stem and progenitor cell; ery, erythroid; GMP, granulocyte macrophage progenitor; ProMono, promonocyte; Mono, monocyte; ncMono, non-classical monocyte; cDC, conventional dendritic cell; pDC, plasmacytoid dendritic cell; Pro-B, pro-B cell; Pre-B, pre-B cell; CD8 term exh, CD8 terminally exhausted; NKT, natural killer T cell; NK, natural killer cell. **b**, UMAP representation of the density of cells from marrow samples of patients with BPDCN ($n = 11$) projected by transcriptional similarity to the cell types defined in **a** and coloured by two-dimensional kernel density estimation. The samples include those without known involvement by BPDCN cells (top; $n = 5$ patients, $n = 36{,}018$ cells) and those with involvement (bottom; $n = 6$ patients, $n = 30{,}582$ cells). BM, bone marrow. **c**, The XV-seq procedure. Mutations identified by DNA sequencing were selected for enrichment on the basis of their detection in matched scRNA-seq data. **d**, UMAP representation of the XV-seq results for enrichment of 16 founder mutations from 5 patients with BPDCN without known marrow

involvement. Cells are projected onto the clusters defined in **a** and coloured according to whether mutant (red; $n = 1,204$ cells) or wild-type (grey; $n = 10,245$ cells) transcripts were detected. Cells without mutant or wild-type calls are not shown. **e**, The expression of BPDCN signature genes (rows; $n = 45$) in cells classified as pDCs from six healthy donors (left; $n = 203$ cells) and six marrow samples from patients with BPDCN involvement (right, $n = 14,209$) (top). Malignant cells are downsampled to 30 cells per sample with genotyping information. The annotation bars (top) indicate the sample identifiers and BPDCN signature scores. Bottom, founder and progression mutations detected by XV-seq. **f**, Expression of BPDCN signature genes (top) and XV-seq mutations (bottom) in cells classified as pDCs from patients without known marrow involvement, as in **e**. Premalignant pDCs (left; $n = 91$ cells with genotyping information) show low signature scores and founder mutations exclusively. Occult BPDCN cells (right; $n = 23$) show high signature scores and a mix of founder and progression mutations. Pt, patient; Dx, diagnosis; Rem, remission; Rel, relapse; WT, wild-type. The diagram in **c** was created using BioRender.

To annotate cells types, we generated a single-cell classifier based on healthy donors using the random-forest machine learning algorithm[26] (Extended Data Fig. 5a). Predicted cell-type proportions in samples without BPDCN involvement were similar to healthy donors, and visualization of marker genes along the myeloid and erythroid differentiation hierarchies did not reveal major abnormalities, consistent with clinical evaluation (Fig. 2b and Extended Data Fig. 5b–d). In samples with known marrow involvement, cells classified as plasmacytoid dendritic cells (pDCs) were highly over-represented (5.3–90.5%, $P = 0.016$, Student's t-test), and application of a published BPDCN gene signature confirmed their identity as malignant BPDCN cells[27] (Fig. 2b and Extended Data Fig. 5c–h). This supports our single-cell approach for the detection of normal and malignant haematopoietic populations from BPDCN marrows.

We next sought to map mutations identified by phylogenomics onto haematopoietic differentiation hierarchies. We prioritized driver mutations and those with high expression by scRNA-seq (eXpressed variant sequencing, XV-seq), reasoning that these would be more efficiently genotyped. In total, we enriched 40 mutations from 11 single-cell libraries using our XV-seq approach and achieved efficient detection for many targets

(for example, *RAB9A*, *CDKN2A*, and *RPS24* in 37%, 20% and 99% of cells, respectively; Fig. 2c, Methods, Extended Data Fig. 6a–e and Supplementary Table 4). The XV-seq findings were in agreement with tumour phylogenies inferred from orthogonal assays, and donor/host annotations in samples after stem-cell transplantation (Extended Data Fig. 6f–h).

To evaluate the differentiation propensities of premalignant clones, we investigated 16 founder mutations in uninvolved marrows from 5 patients and detected a total of 10,245 wild-type and 1,204 mutated cells (Fig. 2d). Mutations were identified across progenitor, erythroid/myeloid and lymphocyte populations, with *TET2* mutations showing erythroid/myeloid bias in 4 out of 5 cases ($P < 0.05$; Fig. 2d and Extended Data Fig. 6i). This confirms the origin of founder mutations in multipotent progenitors rather than lineage-committed pDCs.

Resolving premalignant pDCs and BPDCN

Our analyses suggest a model in which clonal (premalignant) haematopoietic precursors give rise to a spectrum of differentiating blood cell populations, including pDCs. These pDC-committed cells then acquire additional mutations during malignant transformation to BPDCN. To test this, we analysed founder and progression mutations in pDCs from healthy donors and patient marrow samples with and without known involvement by BPDCN.

First, we analysed cells classified as pDCs on the basis of our random-forest approach (203 from healthy donors, 518 from patient samples without marrow involvement and 13,709 from patient samples with marrow involvement). pDCs from healthy donors and patients without marrow involvement clustered together, whereas putative malignant BPDCN cells from patients with marrow involvement clustered separately (Extended Data Fig. 7a). We next generated a BPDCN transcriptional signature by comparing BPDCN cells from cases with marrow involvement to pDCs from healthy donors and patients without marrow involvement (Extended Data Fig. 7b and Supplementary Table 3c). This signature comprised 45 upregulated genes, including *BCL2*, *IGLL1*, *TCL1A* and *HES6*, and showed

strong correlation with a published BPDCN signature[27] (Pearson's $r = 0.89$; Extended Data Fig. 7c).

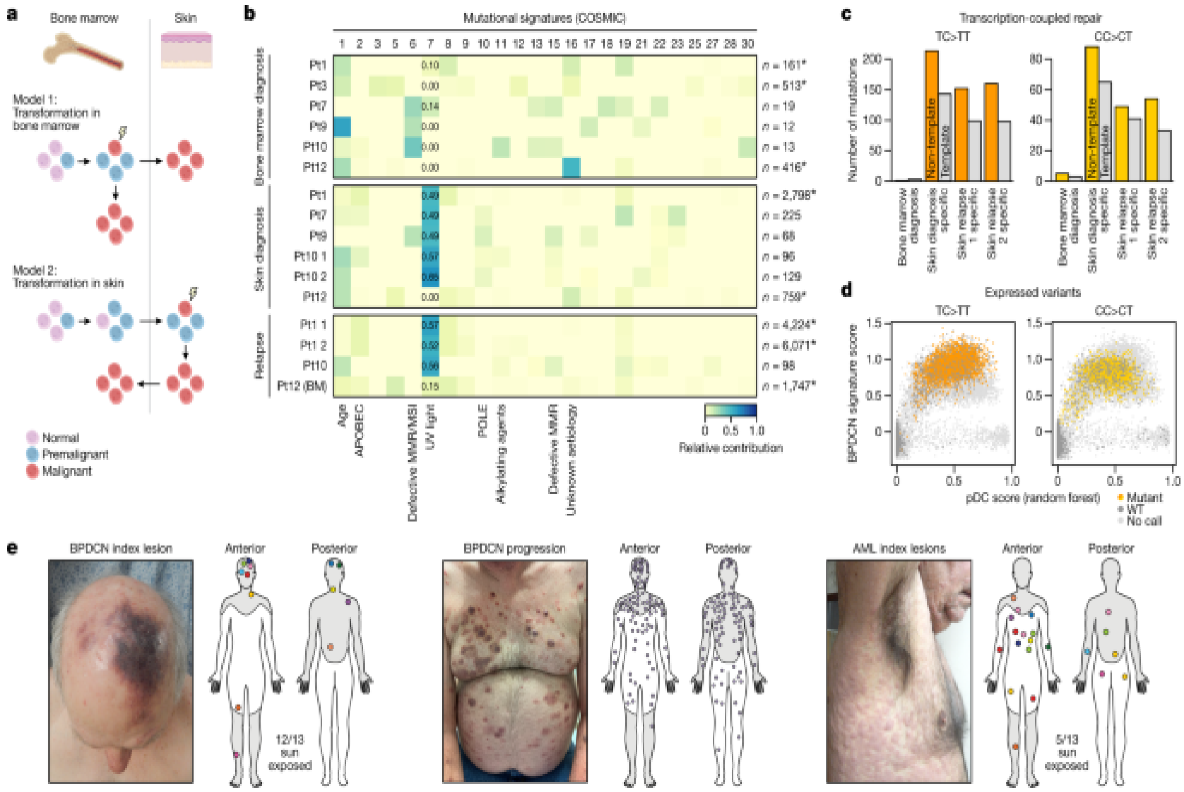
Next, we used single-cell transcriptional and mutation analysis to characterize the stepwise changes leading to pDC transformation. Samples with marrow involvement contained pDCs with high BPDCN signature scores and both founder and progression mutations, consistent with classification as malignant BPDCN cells (Fig. 2e and Extended Data Fig. 7d–h). By contrast, samples without known marrow involvement contained pDCs with low BPDCN signature scores and founder mutations only, probably representing premalignant pDCs or pDC-like cells ($n = 495$ cells; Fig. 2f (left)). Compared with healthy pDCs, premalignant pDCs adopted some but not all of the changes in gene expression observed in BPDCN cells, including upregulation of the *TCL1A* oncogene and downregulation of the *CXCR3* and *CXCR4* chemokine receptors involved in homing to bone marrow and other tissues (Extended Data Fig. 8a,b).

Notably, we also identified rare pDCs with high BPDCN signature scores ($n = 23$) in samples without known marrow involvement, including 2 out of 4,593 cells (0.04%) from patient 9; 19 out of 10,106 cells (0.19%) from patient 10; and 2 out of 6,862 (0.03%) from patient 12 (Fig. 2f (right) and Extended Data Fig. 7d). Although the clinical significance of these rare cells requires further investigation, XV-seq revealed progression mutations in 19 out of 23 (82.3%) pDCs with high scores but in only 1 out of 495 other pDCs, verifying their malignant identity ($P = 1.1 \times 10^{-84}$; Extended Data Fig. 8c–e). Thus, integrated expression and genotyping analysis resolves premalignant pDCs and BPDCN cells, including identification of rare malignant cells not detected by routine histopathology.

UV radiation mutagenesis in BPDCN

The presence of rare malignant cells in bone marrow that were occult to clinical evaluation prompted us to consider two models of BPDCN transformation: (1) transformation of premalignant cells in the marrow followed by spread to the skin; or (2) homing of premalignant cells to the skin, followed by transformation and spread back to bone marrow (retrograde dissemination; Fig. 3a).

Fig. 3: UV damage localizes malignant progression of BPDCN to the skin.



a, Two models of clonal progression to malignancy in BPDCN. In model 1 (top), clonal precursors (blue) transform to malignant cells (red) in the bone marrow, and then spread to the skin. In model 2 (bottom), clonal precursors from the marrow (blue) seed the skin, transform to malignant cells (red) and then spread in a retrograde manner back to bone marrow. **b**, Mutational signature analysis of uninvolved patient marrows (top), BPDCN skin tumours at diagnosis (middle) and relapse samples (bottom; all skin tumours except for patient 12 (Pt12), which is marrow). Blue heat indicates the relative contribution of each signature. The total SNVs per sample is indicated on the right. *n* values marked by asterisks indicate samples profiled by WGS; all others represent WES. **c**, UV-associated TC>TT and CC>CT mutations in samples from patient 1, separated according to their presence on the template (transcribed, indicated in grey) or non-template (non-transcribed, indicated in colour) strands of annotated genes. **d**, Single cells ($n = 66{,}600$) from all marrow samples, showing the random-forest pDC prediction score (*x* axis) and the BPDCN signature score (*y* axis). The colours indicate cells with UV-associated progression mutations (TC>TT,

orange; CC>CT, yellow) detected by XV-seq. **e**, The anatomical distribution of BPDCN skin tumours at the time of diagnosis (left) and disease progression (middle). Skin lesions in patients with AML at diagnosis (right). Grey shading indicates areas of chronic or intermittent UV exposure. Representative clinical photos are shown. The diagram in **a** was created using BioRender.

We reasoned that mutational signature analysis of founder and progression mutations could define the anatomical site(s) at which transformation probably occurred. We therefore evaluated our genomic sequencing data (WES and WGS) for 30 mutational signatures defined in the Catalogue of Somatic Mutations in Cancer (COSMIC). We found a marked enrichment for UV-radiation-associated signatures (signature 7) in BPDCN tumours from 4 out of 5 patients (relative contribution, 0.49–0.65) but not in matched (uninvolved) marrow samples (relative contribution, 0–0.14, $P < 0.005$, Wilcoxon rank-sum test; Fig. 3b). We also detected UV signatures in 10 out of 21 (47.6%) samples from three additional BPDCN cohorts[16,17,28] (using a stringent threshold of 0.2; Extended Data Fig. 9a).

We next examined the nucleotide changes that drive these UV signatures. UV-induced DNA damage is characterized by C>T transitions in dipyrimidine contexts, including UV-associated TC>TT and CC>CT mutations, and UV-specific CC>TT dinucleotide mutations[29]. UV-associated mutations were enriched in BPDCN skin tumours (43.1–55.6% TC>TT, 13.3–37.7% CC>CT; $n = 5$ from patients 1, 7, 9 and 10) but not in matched (uninvolved) marrow samples (5.3–15.4% TC>TT, $P < 0.001$, Wilcoxon rank-sum test; Extended Data Fig. 9b,c). Moreover, UV-specific CC>TT dinucleotide mutations were exclusively detected in BPDCN tumour samples but not uninvolved marrows (Extended Data Fig. 9d). Finally, UV-associated TC>TT and CC>CT mutations were enriched on the non-template strand of coding genes (60.7% and 57.9%, $P < 0.005$, binomial test; Fig. 3c). This probably reflects transcription-coupled repair of mutations on the template strand—another hallmark of UV damage[30,31].

In previous studies, UV mutations in sun-exposed skin from healthy donors have been identified at low a VAF, reflecting background damage to skin cells that do not show clonal expansion beyond minute foci[6]. By contrast,

UV mutations in BPDCN were detected at a high VAF, suggesting their presence as clonally expanded alterations within malignant cells. To validate this, we revisited our scRNA-seq with genotyping data (XV-seq) and assessed the distribution of UV mutations. In total, we identified 4,263 cells in the bone marrow with UV mutations. Among these, 4,100 (96.2%) were classified as malignant BPDCN cells, including the population of rare (occult) BPDCN cells ($n = 19$) from patient 10 described above (Fig. 3d and Extended Data Fig. 9e). The presence of UV-associated mutations in BPDCN cells in the marrow is consistent with their previous transit through the skin.

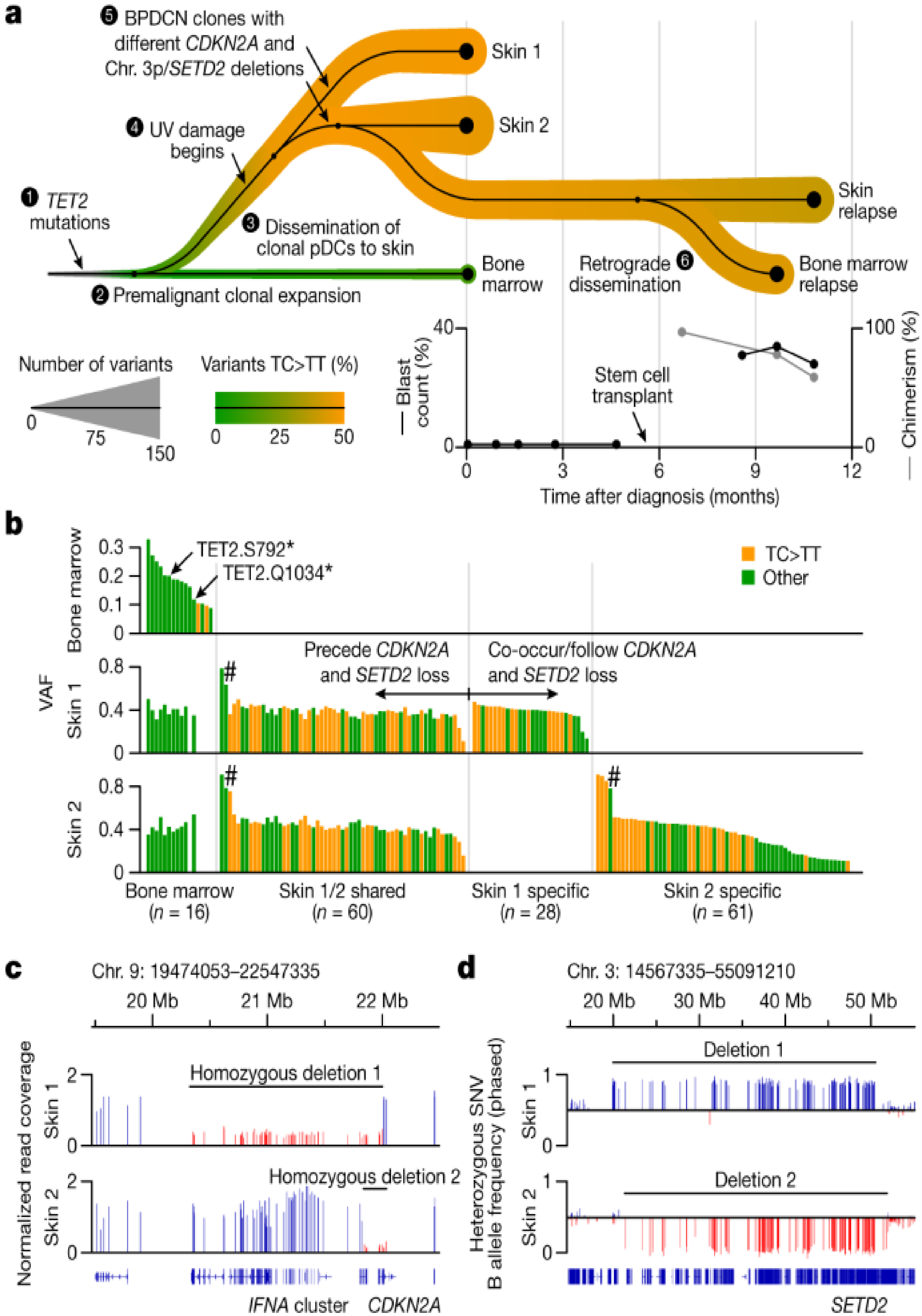
On the basis of these findings, we further hypothesized that BPDCN tumours at the time of initial clinical presentation would localize to sun-exposed rather than sun-protected skin. Indeed, index skin tumours from our cohort were typically solitary lesions at sun-exposed sites, including head and neck, upper back and central chest ($n = 13$; Fig. 3e (left)). By contrast, skin lesions from patients with BPDCN at disease progression (Fig. 3e (middle)) or AML patients at diagnosis tended to be disseminated and involve both sun-exposed and sun-protected sites (Fig. 3e (right); 12 out of 13 initial BPDCN versus 5 out of 13 AML skin lesions in sun-exposed sites; $P = 0.0112$, Fisher's exact test). Together, these findings indicate that UV exposure is a recurrent feature of BPDCN.

UV radiation damage before BPDCN transformation

We next examined the temporal order of UV damage during BPDCN development. Specifically, we examined whether UV exposure begins before or after the occurrence of progression mutations associated with malignant transformation. Accordingly, we overlaid UV mutations onto a detailed tumour phylogeny for patient 10, which was constructed from multiple samples and clinical timepoints. Skin tumours taken from different anatomical sites at diagnosis contained two distinct malignant clones that arose from a common *TET2*-mutated precursor in the marrow (Figs. 1c and 4a,b). Progression mutations in these clones included parallel *CDKN2A* deletions with unique breakpoints (Fig. 4c), and loss of different chromosome 3p alleles in regions containing the tumour suppressor *SETD2*

(Fig. 4d and Extended Data Fig. 10a), indicating convergent evolution during malignant transformation. The majority of UV-associated TC>TT mutations (32 out of 60 alterations, 53.3%) were shared between these two malignant clones (Figs. 3c and 4a,b). This indicates that UV damage began before the loss of *CDKN2A* and *SETD2*.

Fig. 4: UV damage begins before malignant transformation.



a
❺ BPDCN clones with different CDKN2A and Chr. 3p/SETD2 deletions
❹ UV damage begins
❶ TET2 mutations
❸ Dissemination of clonal pDCs to skin
❷ Premalignant clonal expansion
Skin 1
Skin 2
Skin relapse
Bone marrow
Retrograde dissemination ❻
Bone marrow relapse
Number of variants
0
75
150
Variants TC>TT (%)
0
25
50
Blast count (%)
Chimerism (%)
Stem cell transplant
Time after diagnosis (months)
b
Bone marrow
Skin 1
Skin 2
VAF
TET2.S792*
TET2.Q1034*
TC>TT
Other
Precede CDKN2A and SETD2 loss
Co-occur/follow CDKN2A and SETD2 loss
Bone marrow (n = 16)
Skin 1/2 shared (n = 60)
Skin 1 specific (n = 28)
Skin 2 specific (n = 61)
c
Chr. 9: 19474053–22547335
20 Mb
21 Mb
22 Mb
Normalized read coverage
Homozygous deletion 1
Homozygous deletion 2
IFNA cluster
CDKN2A
d
Chr. 3: 14567335–55091210
20 Mb
30 Mb
40 Mb
50 Mb
Heterozygous SNV B allele frequency (phased)
Deletion 1
Deletion 2
SETD2

a, Subway plot showing clonal dynamics, clinical features and the disease course of patient 10. Samples profiled by WES (*n* = 5) are indicated by black dots and connected according to phylogenomic relationships. The line width indicates the total number of detected mutations in each sample, and the colour indicates the percentage of UV-associated TC>TT mutations from green (0%) to orange (50%). The plots at the bottom show the bone marrow blast count (left *y* axis, black lines) from pathology assessment, and donor chimerism (right *y* axis, grey line) after allogeneic stem-cell transplant. **b**, VAFs (*y* axis) of somatic mutations (*x* axis) detected in uninvolved bone marrow (top) and two BPDCN skin tumours from distinct anatomical sites (middle and bottom) in patient 10 at diagnosis, as in **a**. Mutations are grouped according to the sample in which they were first detected (left, bone marrow; middle left, shared skin 1 and 2; middle right, unique to skin 1; right, unique to skin 2). UV-associated TC>TT mutations are indicated in orange, and other mutations are indicated in green. The hashes indicate mutations of which VAFs are affected by copy-number alterations or location on chromosome X. **c**, Normalized read coverage on chromosome 9 for the two BPDCN skin tumours presented in **a** and **b**. Separate homozygous deletions affecting the *CDKN2A* tumour suppressor gene are indicated, with vertical bars showing the read coverage in the deleted (red bars) and non-deleted (blue bars) regions. **d**, Phased B allele frequencies (*y* axis) of heterozygous SNVs on chromosome 3 for the two BPDCN skin tumours presented in **a**–**c**. The vertical bars indicate the allele frequencies in a region containing the *SETD2* tumour suppressor gene. Blue bars, A allele lost; red bars, B allele lost.

To further validate this finding, we compared the VAFs of founder and UV-associated progression mutations in patient 10. In the bone marrow, founder mutations (*n* = 16) showed a wide range of VAFs (coefficient of variation, 0.38), consistent with their gradual acquisition over time (Fig. 4b). By contrast, mutations common to marrow and both skin samples showed a narrow distribution of VAFs (0.13–0.15), consistent with their acquisition in a clonal precursor before transformation. Similar ranges of VAFs were observed for founder mutations in bone marrow (0.31–0.46) and skin tumour samples (0.09–0.10) from patients 7 and 9 (Extended Data Fig. 10b).

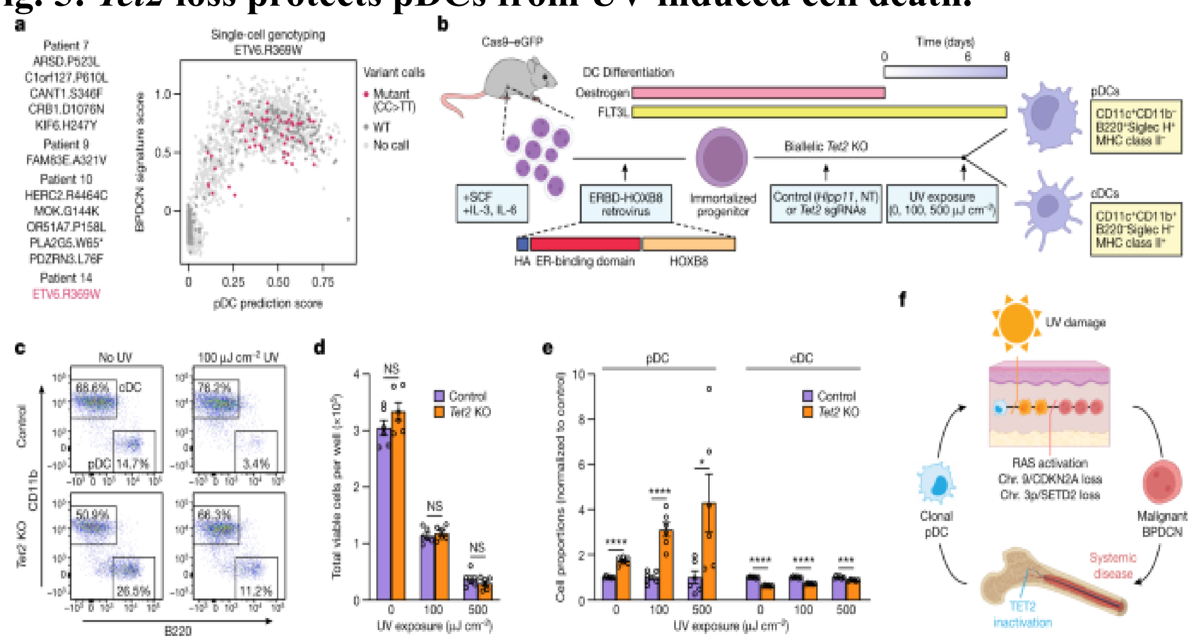
Together, these data indicate that UV damage can precede malignant transformation, and further nominate a clonal (premalignant) pDC or pDC-

like cell in the skin as the precursor of BPDCN in many cases (Fig. 3a (model 2)). Moreover, our data indicate that transformed BPDCN cells in the skin can then spread to other anatomical sites, including through retrograde dissemination to previously uninvolved bone marrow.

Loss of *Tet2* protects pDCs from UV effects

We next evaluated the role of UV damage in malignant transformation and disease evolution. Initially, we reasoned that UV-induced DNA damage would trigger malignant transformation in pDCs through disruption of tumour suppressors or activation of oncogenes. However, as has been demonstrated for skin cancer, assigning UV damage as the causal mechanism for isolated mutations is challenging, unless they happen to involve UV-specific CC>TT substitutions[32]. Among 12 such CC>TT mutations identified in our cohort, one affected a known leukaemia driver gene (*ETV6* p.R369W, patient 14; Fig. 5a (left)) and XV-seq confirmed its presence exclusively within BPDCN cells (Fig. 5a (right) and Extended Data Fig. 10c). Nonetheless, definitive UV-induced oncogenic events could not be identified for the remaining cases in our cohort.

Fig. 5: *Tet2* loss protects pDCs from UV-induced cell death.



a, All detected UV-specific (CC > TT) gene mutations, including ETV6.R369W (red) in patient 14 (left). Right, XV-seq analysis of the bone marrow from patient 14 ($n = 7,374$ cells) showing the random-forest pDC prediction (*x* axis) and BPDCN signature (*y* axis) scores. The colours indicate cells with ETV6.R369W mutant calls (red), wild-type transcripts only (dark grey) and no variant calls (light grey). **b**, Ex vivo differentiation of dendritic cells from mouse bone marrow. Transduction of oestrogen-responsive HOXB8 generates progenitors capable of stable propagation and gene editing in vitro. Oestrogen withdrawal triggers differentiation over 6–8 days into pDCs and cDCs. UV exposure was performed on day 6, and cells were further differentiated until day 8. ER, estrogen receptor. **c**, Representative flow cytometry analysis of cDC ($CD11b^{+}B220^{−}$) and pDC ($CD11b^{−}B220^{+}$) populations in the control, *Tet2*-knockout and UV-exposed conditions. Gating was performed on viable $CD11c^{+}$ cells, as in Extended Data Fig. 10d. **d**, Viable cells (*y* axis) in control and *Tet2*-knockout cells on day 8 after UV exposure on day 6. **e**, The proportion of viable cells (*y* axis) classified as pDCs or cDCs in control (purple) or *Tet2*-knockout (orange) conditions at the indicated UV doses. Data are normalized to the 0 UV condition. **f**, Proposed model for BPDCN development in UV-associated cases. Clonal (premalignant) pDCs/pDC-like precursors arise in the marrow and seed the skin. These cells are then exposed to UV, undergo clonal selection and acquire additional mutations during malignant transformation. Malignant cells then spread systemically, including through retrograde dissemination back to the bone marrow. For **d** and **e**, data are mean ± s.e.m., and include two control and two *Tet2* gRNAs performed in triplicate, representative of two independent experiments. Statistical analysis in **d** and **e** was performed using two-sided Student's *t*-tests; $***P < 0.001$. The diagrams in **b** and **f** were created using BioRender.

Source Data

We next considered another hypothesis whereby UV exposure might provide a selective pressure for premalignant pDCs. To test this, we applied an ex vivo culture system of primary mouse bone marrow progenitors transduced with an oestrogen-responsive form of *HOXB8*[33] (Fig. 5b). In the presence of oestrogen and FLT3 ligand, this system enables stable propagation and genome editing of myeloid progenitor cells in vitro. Oestrogen withdrawal

can then be used to trigger synchronized differentiation over 6–8 days into mature pDCs and conventional DCs (cDCs)[33] (Extended Data Fig. 10d,e).

We exposed these cultures to UV radiation on day 6 of differentiation after oestrogen withdrawal. By titrating UV exposure down to a low level, we identified a dose range (0, 100, 500 μJ cm^{-2}) that induced a reproducible gradient of cell death over a 24–48 h period[34] (Fig. 5c,d). We next investigated whether premalignant pDCs were more resistant to UV-induced cell death compared with normal pDCs or cDCs. We used CRISPR–Cas9 to generate HOXB8 progenitors with biallelic *Tet2* mutations—the most common premalignant alteration in BPDCN—and then differentiated these cells into mature pDCs and cDCs (Extended Data Fig. 10f). In the absence of UV, *Tet2* knockout enhanced the proportion of pDCs (1.72-fold, $P < 0.0001$), consistent with a role for TET2 in DC differentiation[17,35] (Fig. 5c). After UV exposure (100, 500 μJ cm^{-2}), *Tet2* knockout caused a further increase in the proportion of surviving pDCs, but had no protective effect on cDCs (Fig. 5e). This suggests a tumour suppressor role for TET2 in UV-exposed pDCs, and may explain the strong association between *TET2* inactivation, skin localization and UV-associated mutations in BPDCN.

Discussion

Here we studied the development of BPDCN, an unusual acute leukaemia that originates from clonal precursors in the bone marrow but often presents with malignant cells isolated to the skin. Using tumour phylogenomics and single-cell profiling, we found that premalignant clones in the bone marrow give rise to BPDCN tumours in the skin. BPDCN cells are distinguished from clonal precursors by numerous UV-associated mutations that are absent from other blood-cell lineages. We found that UV-associated mutations can begin accumulating before malignant transformation and that BPDCN skin tumours first appear at sun-exposed sites. Taken together, this implicates UV exposure of premalignant pDCs or committed precursors in the skin during tumour pathogenesis. Moreover, we show that malignant cells from the skin can spread through retrograde dissemination to previously uninvolved bone marrow, and integrate single-cell gene expression and mutation analysis to identify rare circulating UV-damaged BPDCN cells (Fig. 5f).

Although pDCs are not present at high numbers in normal skin, they are recruited during antiviral, autoimmune and wound-healing responses after skin injury or UV damage[36,37,38]. pDCs recruited to the skin can be eliminated by UV phototherapy, which may contribute to the benefit of this treatment for inflammatory dermatoses[36,39]. Consistent with this, we show that normal pDCs are highly sensitive to UV-induced cell death, whereas *Tet2*-mutated pDCs are relatively protected. These data suggest that biallelic inactivation of *TET2*, a highly recurrent premalignant alteration in BPDCN, facilitates pDC expansion and mutagenesis. We propose that UV exposure contributes to the anatomical localization and selective pressures faced by clonal pDCs or precursors during the early stages of BPDCN development in the skin, and may explain the high fraction of *TET2* alterations observed in this malignancy. Areas for future research include determining how TET2 affects the DNA damage response, and whether UV mutations create neoantigens that could sensitize to immunotherapy.

Finally, our findings demonstrate how exposure of circulating (clonal) immune cells to tissue-specific environmental exposures can shape mutational evolution and disease progression. Although BPDCN is relatively rare, UV signatures have also been observed in normal and malignant T cell populations and in a subset of paediatric B lymphoblastic leukaemia/lymphoma[40,41,42]. Together, these data support a model for clonal selection and premalignant evolution in the blood system that can be influenced by more than one tissue environment. More broadly, the role of tissue-specific mutational processes in shaping the evolution of clonal disorders across anatomical sites warrants further investigation.

Methods

Patient samples

Patients with BPDCN seen at the Dana-Farber Cancer Institute provided informed consent to an IRB-approved research protocol permitting tissue collection and sequencing analysis. The demographic characteristics of the patient cohort are provided in Supplementary Table 1a. Healthy control participants for single-cell sequencing consented to IRB-approved research protocols from Brigham and Women's Hospital or Lonza Bioscience that

cover all of the study procedures; demographics are provided in Supplementary Table 3a.

Histopathology

Histological processing and immunohistochemical staining of patient bone marrow and skin tumour biopsies was performed according to routine clinical procedures in the Department of Pathology at the Brigham and Women's Hospital, as previously described[26,43,44]. Results are included in Supplementary Table 1a.

Targeted DNA sequencing and analysis

Targeted sequencing of fresh bone marrow samples using a 95-gene leukaemia panel (Rapid Haem Panel, $n = 23$ samples) and formalin-fixed, paraffin-embedded archival skin tumour samples using a 282-gene pan-cancer panel (Oncopanel, $n = 9$ samples) was performed for genes recurrently altered in myeloid malignancies and BPDCN[45,46,47,48]. Mutation calls were manually inspected and verified in sequence alignment files for each sample. Combined mutation calls and VAFs across all samples are provided in Supplementary Table 2a. For the patient 2 bone marrow sample, we observed a lower read coverage for amplicons covering the *ASXL1* gene relative to other samples and controls (67%, corresponding to a VAF of 0.33, located on chromosome band 20q11) and, accordingly, included this copy-number alteration in Fig. 1b.

WES analysis

WES analysis of cryopreserved bone marrow, skin tumour and germline samples (uninvolved skin for patient 7, bone-marrow-derived fibroblasts for patients 9 and 10) was performed using the Illumina HiSeq 4000 (2 × 150 bp, patient 7) and the BGISEQ-500 (2 × 100 bp, patients 9 and 10) platforms, as previously described[49]. An overview of all of the profiled samples is provided in Supplementary Table 1b. Sequencing data for a total of 12 samples were mapped to the human genome reference (hg19; https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_000001405.13/) using BWA (v.0.7.15)[50]. The resulting BAM files were further analysed and

recalibrated with Picard (v.2.5.0)[51] and the GATK toolkit (v.4.0.0.0)[52]. Somatic mutations were identified using Mutect2[53] by comparing to patients' germline variations. Initial calls were filtered by estimated cross-sample contamination and artifacts related to orientation bias. Calls were then merged for each patient and further filtered by removing calls that showed a VAF lower than 0.1 in all samples, with the exception of mutations that were also identified by targeted sequencing (that is, mutations in *TET2* in patient 10). We also excluded calls that showed a germline VAF that was greater than one-fifth of the highest VAF of other samples from the same patient, and mutations that were detected across multiple patients (with the exception of hotspot mutations in *ASXL1* in patients 7 and 10). Resulting high-confidence mutation calls for each sample are provided in Supplementary Table 2b–d. For the patient 10 relapse bone marrow sample, which was collected after the patient received a stem cell transplant, we were unable to define sample-specific mutations owing to the high proportion of donor DNA. However, we could quantify mutations in this sample identified in other samples from the same patient.

Combined mutation calls were the basis for the inference of tumour phylogenies (Figs. 1c and 4a) and the inference of putative clonal architectures in patient bone marrow samples (Fig. 1d). Mutation calls were further analysed for mutational signatures defined in the COSMIC database[54] by applying the R package MutationalPatterns[55]. The relative contribution of 30 different mutational signatures was calculated and scores for UV-light-associated signature 7 are indicated (Fig. 3b). For copy-number analysis, SNVs were jointly identified for all samples from each patient using bcftools (v.1.10.2; commands mpileup and call). The B (minor) allele frequency and read coverage (relative to germline samples) for each SNV was used to infer copy-number alterations (Extended Data Figs. 2 and 3). Public datasets (Extended Data Fig. 8a) were analysed from provided mutation calls, with the exception of data from ref. [28], which were processed from raw sequencing reads (Sequence Read Archive: SRP301976) and analysed using Mutect2.

WGS analysis

WGS analysis of germline, bone marrow and skin tumour samples at diagnosis and relapse (13 samples from patients 1, 3 and 12) was performed using the BGI DNBSEQ platform (2 × 100 bp). Bone marrows were profiled from cryopreserved samples, skin tumours were profiled from formalin-fixed paraffin-embedded (FFPE) samples and germline samples were profiled from both sample types (Supplementary Table 1b). Sequencing data were mapped to the human genome (hg19) reference using BWA (v.0.7.17)[50]. The resulting BAM files for all of the samples from each patient were jointly analysed using Mutect2[53] by comparing to the matched germline sample, supplying both a germline resource (somatic-b37_af-only-gnomad.raw.sites.vcf) and a panel of healthy individuals (somatic-b37_Mutect2-exome-panel.vcf). Variant calls were filtered using GATK[52] (v.4.2.3.0; commands LearnReadOrientationModel and FilterMutectCalls) and annotated using the Funcotator command (funcotator_dataSources.v1.6.20190124s). The resulting calls were further filtered by retaining only variants that had a coverage greater than 20 for cryopreserved samples and greater than 10 for FFPE samples in all samples per patient. Final mutation calls were defined at the latest timepoint per patient: in patient 1, we considered variants detected in both skin relapse 2 and bone marrow relapse (collected after the patient received a stem cell transplant) with a VAF of greater than 0.25. In patient 3, we considered variants detected in the diagnostic bone marrow with a VAF of greater than 0.1. In patient 12, we considered variants detected in bone marrow relapse with a VAF of greater than 0.25. These filtering steps were deemed to be appropriate owing to the lower quality of FFPE-derived samples, and challenges due to high proportions of donor DNA in patients who received a stem cell transplant. Variants were attributed to the first sample in which they were detected with a VAF greater than 0.1, and all subsequent samples.

The resulting mutation calls were visualized in tumour phylogenies (Extended Data Fig. 1d) and were analysed for mutational signatures similar to as described for the WES dataset above (Fig. 3b). Summaries of mutation calls for each sample are provided in Supplementary Table 2e. For copy-number analysis, single-nucleotide variants were jointly identified for all of the samples from each patient using bcftools (v.1.10.2; commands mpileup and call). The B allele frequency and read coverage (relative to germline

samples, not shown) for each SNV was used to infer copy-number alterations (Extended Data Fig. 4).

scRNA-seq

scRNA-seq was performed on cryopreserved bone marrow aspirates. Cells were stored in liquid nitrogen, thawed using standard procedures and viable (propidium iodide negative) cells were sorted on the Sony SH800 flow cytometer. Next, 10,000–15,000 cells were loaded onto a Seq-Well array or 10x Genomics chip. Further processing was performed using the recommended procedures for the Seq-Well S^3 (http://shaleklab.com/resource/seq-well/)[56] or the 10x Genomics 3′ v3/v3.1 chemistry. Seq-Well S^3 libraries were sequenced on the NextSeq system (20 + 8 + 8 + 57 cycles) and 10x libraries were sequenced on the NovaSeq system (28 + 8 + 91 cycles for single-index or 28 + 10 + 10 + 75 cycles for dual index). Some of the data were previously reported[57,58] (Supplementary Table 3a). Serial samples from the same patient were loaded onto separate sequencing runs to avoid erroneous assignment of reads by index swapping between diagnosis/remission/relapse samples (this is particularly relevant for the identification of rare malignant cells).

XV-seq

We developed an improved method for targeted enrichment of genetic variants from scRNA-seq libraries that is compatible with Seq-Well S^3 and 10x 3′ gene expression platforms. Compared to previous methods by us and others[26,59], we incorporated a number of computational and experimental steps for increased sensitivity and specificity: (1) we considered all mutations detected by WES, including synonymous mutations and mutations affecting untranslated regions (UTR). These mutations do not result in changes in the protein sequence, but can be used to infer clonal relationships. (2) We quantified detection of these mutations in the regular scRNA-seq data before enrichment. For example, of the 186 mutations detected across samples for patient 10, only 16 (8.6%) were detected in at least one transcript (Supplementary Table 2d). We found that detection in the regular scRNA-seq data is a good predictor of enrichment efficiency (Extended Data Fig. 6c). (3) We specifically considered loci of which only a

single allele is present in the genomes of healthy and/or malignant cells. For these mutations, detection of the wild-type allele is as informative as the presence of the mutant allele (that is, if the wild-type is detected, the mutant must be absent; for heterozygous mutations, the mutant could remain undetected). In our dataset, this included a mutation in the X-chromosomal gene *RAB9A* (in a male patient), a focal deletion of *CDKN2A/B*, which occurred in cells already carrying loss of chromosome 9, and 3′ UTR mutations in *SETX* and *SMARCC1*, which also occurred in cells with loss of the other allele on chromosome 9 and chromosome 3. (4) Finally, we incorporated technical optimizations such as inclusion of dual indices, as outlined below.

XV-seq for Seq-Well S^3

Compared with single-cell genotyping of Seq-Well S^3 libraries that we previously reported[26], we adjusted primer designs to generate dual-indexed libraries. This increases the confidence that reads are assigned to the correct library, particularly when using Illumina instruments with patterned flow cells. We first designed biotinylated mutation-specific primers to detect each of the known mutations in a given sample (Supplementary Table 4a). As a starting material, we used amplified cDNA from the Seq-Well S^3 protocol (also known as whole-transcriptome-amplified material). We then set up a biotin-PCR reaction to add a biotin tag and Nextera adapter to our gene of interest while retaining the unique molecular identifier (UMI) and cell barcode, as follows. We created a mixture containing a standard reverse primer at 3 µM (SMART-AC), and mutation-specific primers at a combined concentration of 3 µM. To prepare the template for the biotin-PCR reaction, we pooled and diluted whole-transcriptome-amplified products from the same sample and timepoint to 10 ng in a total volume of 10 µl. We next added 2.5 µl of primer mix (final concentration, 0.3 µM) and 12.5 µl of 2× KAPA HiFi Hotstart Readymix (Roche, NC0465187) to the template. We performed PCR using the following conditions: initial denaturation at 95 °C for 3 min; followed by 12 cycles of 90 °C for 20 s, 65 °C for 15 s and 72 °C for 3 min; and final extension at 72 °C for 5 min. After amplification, we purified the PCR product with 0.7× AMPure XP beads and captured biotinylated fragments using Streptavidin beads.

To add Illumina adapters, dual-indexed barcodes and a custom read primer binding sequence to the fragments, we performed a second PCR using the Streptavidin-bound product as a template (23 µl), with 2 µl of a 5 µM primer mix (N70D primers, Supplementary Table 4b) and 25 µl PFU Ultra II HS 2× Master Mix (Agilent, 600850). The parameters used for the second biotin-PCR were as follows: initial denaturation at 95 °C for 2 min; then 4 cycles of 95 °C for 20 s, 65 °C for 20 s and 72 °C for 2 min; followed by 10 cycles of 95 °C for 20 s and 72 °C for 2 min and 20 s; and then final extension at 72 °C for 5 min. After the second PCR, we magnetized the Streptavidin beads and saved/purified DNA from the supernatant with 0.7× AMPure XP beads. After eluting in 20 µl of TE, we magnetized the beads and saved the supernatant for sequencing on the Illumina NextSeq system.

XV-seq for 10x

We adjusted the Genotyping of Transcriptomes[59] protocol by (1) omitting staggered handles on gene-specific primers and (2) incorporating dual 10 bp library indices, which minimizes the chance of barcode swapping and ensures compatibility with 10x Genomics scRNA-seq v3.1 libraries and Illumina v1.0 and v1.5 chemistry. The starting material for transcript genotyping were the full-length cDNA libraries generated according to the 10x Genomics 3' v3 or v3.1 scRNA-seq protocol. If cDNA quantities were limited, we performed a full-length cDNA PCR amplification using generic primers that bind to all transcripts (primers: PartialRead1 and PartialTSO; Supplementary Table 4c). The pre-enrichment PCR was set up by mixing 10 ng of cDNA template, forward and reverse primers at 0.3 µM each, 2× Kapa HiFi HotStart ReadyMix and H_2O up to 50 µl. The PCR was performed under the following conditions: initial denaturation at 95 °C for 3 min; followed by 6 cycles of 98 °C for 20 s, 67 °C for 15 s, 72 °C for 3 min; and a final extension of 72 °C for 3 min. After amplification, we purified the PCR product with 0.6× AMPure XP beads (Beckman Coulter Life Sciences, A63881).

The enrichment for loci of interest consists of two PCR reactions. For PCR1, to enrich for loci of interest (determined by targeted or exome sequencing), we designed primers to amplify specific regions (Supplementary Table 4a). We downloaded the transcript sequence in Geneious Prime 2020, and

annotated the mutation of interest. We designed mutation-specific primers within 50 bases upstream of the mutation site (so that the mutation site would be captured in read 2 of the sequencing data). To add a read 2 sequence to this mutation-specific primer, we appended CACCCGAGAATTCCA at the 5′ end. PCR1 was performed using these mutation-specific primers and a generic forward primer (PartialRead1; Supplementary Table 4c). We mixed up to six mutation-specific primers per PCR1 reaction, as long as they targeted different transcripts. We prepared PCR1 reactions as follows: 100 ng cDNA was added to 0.25 μM forward primer and 0.25 μM mutation-specific primer(s), 20 μl 2× Kapa HiFi HotStart ReadyMix and H_2O up to 40 μl. The PCR was performed under the following conditions: a denaturation step at 95 °C for 3 min; followed by 10 cycles of 98 °C for 20 s, 67 °C for 15 s, 72 °C for 3 min; and a final extension 72 °C for 3 min. After amplification, we purified the PCR product with 1× AMPure XP beads. We next performed PCR2 to generate indexed libraries compatible with the Illumina NextSeq and NovaSeq machines.

For PCR2, we used a P5 sequence followed by a 10 bp index barcode and a read 1 sequence as a forward primer (XV-P5-i5-BCXX) and a P7 sequence followed by a 10 bp index barcode and a read 2 sequence as a reverse primer (XV-P7-i7-BCXX; Supplementary Table 4c). The PCR was set up as follows: 18 μl of the PCR1 product was added to 2 μl primers (0.25 μM each) and 20 μl Kapa HiFi HotStart ReadyMix. The PCR was performed under the following conditions: 95 °C for 3 min; followed by 6 cycles of 98 °C for 20 s, 67 °C for 15 s, 72 °C for 3 min; and a final extension 72 °C for 3 min. After amplification the PCR product was purified with 1× AMPure XP beads. Elution in 20 μl buffer TE typically yielded 5–50 ng μl^{-1} with an average size of 300–1,500 bp, which was pooled for sequencing on the Illumina NextSeq or NovaSeq instruments with the goal of generating 10 million reads per library.

scRNA-seq computation analysis

Data from the Seq-Well protocol (healthy donor 6 and patient 9) were processed as described previously[26]. In brief, demultiplexed fastq files were processed to maintain only cell barcodes with 100 reads and to append the cell barcode and UMI, derived from read 1, to the read identifier of read 2.

The hg38 reference genome and annotations were downloaded from Ensembl (release 99), extended with *RNA18S* and *RNA28S* genes from UCSC, and finalized using Cell Ranger mkgtf with the recommended settings and the additional gene biotypes Mt_rRNA and rRNA. We then used STAR (v.2.6.0c) to align processed fastqs to hg38 and created a count matrix. Data from the 10x Genomics 3′ v3 and v3.1 platform (the remaining 15 samples) were processed using Cell Ranger (v.7.0.0) using the default settings and the same hg38 reference. Count matrices from both the Seq-Well and the Cell Ranger pipelines were processed to retain only cells with >2,000 UMIs, >1,000 genes and <20% mitochondrial alignments. From the count matrix, we removed mitochondrial genes (^MT-*), genes of the biotype rRNA (defined in the reference gtf file) and *RNA18S*/*RNA28S*. We maintained X- and Y-chromosomal genes, including *ZRSR2* and *IL3RA*.

XV-seq computational analysis

To quantify detection of mutations in regular scRNA-seq libraries, we assessed every mutation detected by exome sequencing in the genome alignments for the respective sample. Mutations were quantified using samtools mpileup. For each base, information for cell barcode and UMI was obtained by setting the --output-extra option, and subsequently collapsed using R and the data.table package. Mutations that were most efficiently detected or that were of special interest were selected for XV-seq enrichment.

For analysis of XV-seq data, fastq files were processed using IronThrone-GoT (v.2.1) using the recommended set-up (https://github.com/landau-lab/IronThrone-GoT)[59]. For patient 9, we used --bclen 12, --umilen 8 and a whitelist of cell barcodes that passed RNA-seq quality controls. For all of the other samples, we used --bclen 16, --umilen 12 and the whitelist 3M-february-2018.txt. For every mutation, we generated custom configuration files to distinguish between wild-type and mutant transcripts by one or several differing bases. If the mutation site was directly adjacent to the primer, the 3′ end of the primer was used as a shared sequence and additional bases were added to the wild-type/mutant sequences, taking into account that IronThrone-GoT allows for 20% of the bases in the analysed reads to be mismatched from the provided sequences. For *MTAP*, five

configuration files were used, one for each of the potential splicing products indicating the *CDKN2A/B* deletion (Extended Data Fig. 6b). IronThrone-GoT jobs were submitted in Linux using the Sun Grid Engine with the options -pe smp 4 -binding linear:4 -l h_vmem=32g -l h_rt=96:00:00. After completion of the IronThrone-GoT run, we processed the generated information (summTable) by plotting the number of wild-type and mutant calls for different sequencing reads of each transcript (UMI). We used only transcripts that were supported by ≥3 reads and with at least threefold more wild-type than mutant calls or vice versa. For MALAT1.n.G3541A in Fig. 2e,f and Extended Data Fig. 8, we reduced the read threshold to 1. In the case of heterozygous mutations, cells in which a wild-type transcript is detected are not necessarily wild-type cells, as the mutated allele may have been missed. In the case of multiple mutations within the same gene (as is observed for *TET2* in BPDCN), transcripts may show a wild-type result at one site while still harbouring a mutation in *cis* at a different position in the same transcript/allele. We added the genotyping information as metadata to Seurat objects with scRNA-seq expression data by joining based on cell barcodes.

To check the accuracy of our single-cell mutation calls, we validated the ASXL1.G642fs mutation in Pt10Dx using two different enrichment primers. This known oncogenic guanine insertion, resulting in ATCGGAGGGGGGGT>ATCGGAGGGGGGGGT, can be challenging to call. We enriched the mutation site from 10,106 high-quality single-cell transcriptomes using two different primers: ASXL1-1886 (CACCCGAGAATTCCA**GTCACCACTGCCATAGAGAGG**) and ASXL1-1898 (CACCCGAGAATTCCA**ATAGAGAGGCGGCCACCA**; the first one is included in Supplementary Table 4a) (transcript-binding sequences are in bold). In the first experiment, we detected mutated *ASXL1* transcripts in nine cells. In the second experiment, we detected mutated transcripts in eight cells. Seven of the cells overlapped between the two attempts, indicating striking concordance. We also detected wild-type *ASXL1* transcripts in 33 and 32 cells in the two experiments, respectively. There was perfect overlap in 32 wild-type cells that were called between the two experiments with different *ASXL1* enrichment primers. The agreement between these experiments, together with the orthogonal targeted DNA sequencing, which identified the same mutation, attests to the reliability of our mutation calls.

Dimensionality reduction and cell type annotation

Count matrices from healthy donors were imported into R (v.4.2.1) using Seurat (v.4.1.1) on a MacBook Pro with an M1 Max chip. Normalization, variable feature identification and data scaling were performed using the Seurat defaults[60]. After principal component analysis, Harmony was used to integrate data from Seq-Well S^3 and 10x v3 3′ scRNA-seq[61]. We then used Harmony reduction to determine clusters and UMAP coordinates. Integration from Seq-Well and 10x platforms generated clusters that were driven by biological (rather than technical) differences between cells (Fig. 2a). Clusters were annotated by expression of canonical marker genes such as *CD34* (progenitors), *CD14* (monocytes), haemoglobin (erythroid), *IRF8* and *TCF4* (pDCs; Supplementary Table 3b). This yielded 21 healthy reference cell types. One cluster (1.04% of healthy donor cells) was classified as doublets on the basis of co-expression of marker genes.

To annotate cell types from the samples of patients with BPDCN, we used the random-forest algorithm[26]. Specifically, we used the R package randomForest (v.4.7-1.1) to generate a classification forest using marker genes of healthy donors (determined by Seurat's FindAllMarkers function); we previously showed that this approach performs similarly to other reference-based classification algorithms[57]. The confusion matrix and fivefold cross-validation both indicated 89.7% accuracy (Extended Data Fig. 5a). We next used the classification forest to assign each cell from the patient samples with prediction/probability scores for each reference cell type (function predict() with randomForest object and type = "prob"; see 3_RandomForest.R at https://github.com/petervangalen/Single-cell_BPDCN/). The reference cell type with the maximum prediction score was used for the patient cell classification. Cells that were classified as doublets (up to 2.34%) were excluded from further analysis. Projection of patient cells onto the UMAP of healthy donor cells was performed by plotting each patient cell at the coordinates of the normal cell with the highest prediction score correlation (Fig. 2b).

Annotation of host and donor single cells

To annotate single cells from the patient 10 relapse bone marrow sample for their origin (this patient received an allogeneic stem cell transplantation prior to relapse), we quantified SNVs specific to the host or donor genome in each single cell. We first identified all SNVs in the relapse bone marrow exome sequencing dataset, which represents a mixture of both genomes (n = 127,916; Extended Data Fig. 3b; see also the copy-number analysis above). For each SNV, we then quantified its B allele frequency in both the germline and relapse bone marrow sample. By applying thresholds on both frequencies, we identified variants that are informative for each genome (Extended Data Fig. 5e). We further removed variants that were located within broad copy-number alterations on chromosomes 3, 6 and 9, as well as on chromosomes X and Y. A total of 56,155 SNVs were identified in this manner, with 5,989 (10.7%) being homozygous in both genomes (that is, host A/A and donor B/B, or host B/B and donor A/A) and therefore informative for both alleles. These SNVs were quantified in the single-cell transcriptome data of the diagnostic and relapse sample using samtools mpileup. For each base, information for cell barcode and UMI was obtained by setting the --output-extra option. We then aggregated coverage for all host- and donor-specific alleles across the genome for each single cell. Cell annotations for the patient 10 relapse bone marrow sample were obtained for cells with a coverage of at least 10 and a donor-specific allele coverage of less than 10% (host cells, n = 4,453) or greater than 90% (donor cells, n = 2,664; illustrated in Extended Data Fig. 5f). A small fraction of cells with donor-specific allele coverage between 10% and 90% potentially reflect cell multiplets and were removed from further analysis. In total, 94.9% of patient 10 relapse cells were annotated for their host/donor origin. As a control, none of the single cells from the patient 10 diagnostic bone marrow sample were classified as donor cells.

BPDCN signature generation and single-cell gene expression analysis

To generate a single-cell transcriptional signature specific for malignant BPDCN cells, we made two groups of cells: (1) cells classified as pDCs from healthy donors and (2) cells classified as pDCs from patients with marrow involvement. Cells classified as pDCs without progression mutations from patients without marrow involvement were also included in

group 1 as they were similar to pDCs from healthy donors (Extended Data Fig. 7a). Cells classified as pDCs with progression mutations from patients without marrow involvement were not included in differential gene expression analysis because they were suspected circulating malignant BPDCN cells based on mutation and gene expression patterns. We randomly selected at most 50 cells per sample, so that the analysis would not be dominated by samples with a high number of pDCs. We then compared the two groups using the Seurat function FindMarkers and selected genes with twofold higher expression in the second group ($\log_2$-transformed fold change > 1) and an adjusted $P < 1 \times 10^{-30}$ (ref. [62]; Wilcoxon Rank Sum test, 45 genes; Extended Data Fig. 7b and Supplementary Table 3c). To identify malignant cells, we scored all 87,011 single-cell transcriptomes for this 45-gene signature using the Seurat function AddModuleScore (Extended Data Fig. 7c,d). We then defined malignant BPDCN cells as all cells classified as pDCs in patients with bone marrow involvement, as well as all cells (regardless of their initial classification) with a BPDCN signature score exceeding 0.5 (Extended Data Fig. 7f–h). To ensure that reclassification of a small proportion of cells as malignant BPDCN cells was justified, we checked marker gene expression, reasoning that the absence of canonical markers would support reclassification. Indeed, reclassified pro-B cells lacked CD19 and reclassified plasma cells lacked CD138 (Extended Data Fig. 7g). Using Seurat objects with scRNA-seq expression data and metadata (including cell type annotations and XV-seq mutation calls joined based on cell barcodes), we performed all downstream single-cell analyses in R with extensive use of the tidyverse[63].

In vitro differentiation of dendritic cells and UV exposure

HOXB8-FL cells were derived as described previously[33] from bone marrow cells of mice constitutively expressing Cas9 (Jackson Laboratory, 026179). HOXB8-FL cells were cultured in RPMI-1640 (Gibco, 11875093) supplemented with 10% FBS (Sigma-Aldrich, F2442), 1% penicillin–streptomycin (Corning, 30002CI), 50 ng μl^{-1} mouse FLT3L (BioLegend, 550706) and 1 μM oestrogen (Sigma-Aldrich, E2758). Cells were resuspended in fresh medium every 2–3 days. For DC differentiation, HOXB8-FL cells were washed once in RPMI-1640, then resuspended in the same medium without oestrogen. For UV exposure on day 6 after-oestrogen

withdrawal, cells were resuspended in PBS and exposed to the indicated doses of UV using an XL-1500 Spectrolinker. Cells were then resuspended in differentiation medium and analysed by flow cytometry on day 8 after oestrogen withdrawal. For CRISPR-mediated knockout of *Tet2*, Cas9-expressing HOXB8-FL cells underwent lentiviral transduction of sgRNA using the pLKO5.sgRNA.EFS.GFP vector (Addgene, 57822). Data were combined from two Tet2-targeting sgRNAs (GAATACTATCCTAGTTCCGAC and GAACAAGCTCTACATCCCGT). For controls, data were combined from non-transduced cells and cells transduced with sgRNA targeting a safe harbour region (ATGTACAACACAAACGAAGT). *Tet2*-sgRNA-induced indels were validated using PCR amplicon next-generation sequencing (Extended Data Fig. 10f). For flow cytometry analysis, differentiated HOXB8-FL cells were incubated for 10 min in Fc block (BD, 553141), then stained for CD11b Alexa Fluor 700 (BioLegend, 101222), CD11c PE/Cyanine7 (BioLegend, 117317), B220 APC/Cyanine7 (BioLegend, 103224), Siglec-H PE (BioLegend, 129605) and MHC-II PerCP/Cyanine5.5 (BioLegend, 107626). cDCs were defined as $CD11c^{+}CD11b^{+}B220$-MHC-II^{+} and pDCs were defined as $CD11c^{+}CD11b^{-}B220^{+}Siglec$-$H^{+}$ (Extended Data Fig. 10d,e). DAPI staining was used to exclude non-viable cells.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The WES/WGS data are available in the dbGaP under accession number phs003228. The single-cell sequencing data and gene expression matrices are available at the Gene Expression Omnibus under accession number GSE227690. Source data are provided with this paper.

Code availability

Single-cell code is shared at GitHub and Zenodo (https://github.com/petervangalen/Single-cell_BPDCN and https://doi.org/10.5281/zenodo.7746255), and Seurat objects are available online (https://vangalenlab.bwh.harvard.edu/resources/).

References

1. Jaiswal, S. et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).

2. Genovese, G. et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).

3. Xie, M. et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).

4. Belizaire, R., Wong, W. J., Robinette, M. L. & Ebert, B. L. Clonal haematopoiesis and dysregulation of the immune system. *Nat. Rev. Immunol.* https://doi.org/10.1038/s41577-023-00843-3 (2023).

5. Garnache-Ottou, F. et al. How should we diagnose and treat blastic plasmacytoid dendritic cell neoplasm patients? *Blood Adv.* **3**, 4238–4251 (2019).

6. Martincorena, I. et al. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).

7. Moore, L. et al. The mutational landscape of normal human endometrial epithelium. *Nature* **580**, 640–646 (2020).

8. Kakiuchi, N. & Ogawa, S. Clonal expansion in non-cancer tissues. *Nat. Rev. Cancer* **21**, 239–256 (2021).

9. Jaiswal, S. et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N. Engl. J. Med.* **377**, 111–121 (2017).

10. Miller, P. G. et al. Association of clonal hematopoiesis with chronic obstructive pulmonary disease. *Blood* **139**, 357–368 (2022).

11. Agrawal, M. et al. TET2-mutant clonal hematopoiesis and risk of gout. *Blood* **140**, 1094–1103 (2022).

12. Khanlari, M. et al. Bone marrow clonal hematopoiesis is highly prevalent in blastic plasmacytoid dendritic cell neoplasm and frequently sharing a clonal origin in elderly patients. *Leukemia* https://doi.org/10.1038/s41375-022-01538-9 (2022).

13. Taylor, J. et al. Multicenter analysis of outcomes in blastic plasmacytoid dendritic cell neoplasm offers a pretargeted therapy benchmark. *Blood* **134**, 678–687 (2019).

14. Swerdlow, S. H. et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues* Revised 4th Edn (IARC, 2017).

15. Suma, S. et al. Blastic plasmacytoid dendritic cell neoplasm arising from clonal hematopoiesis. *Int. J. Hematol.* **108**, 447–451 (2018).

16. Batta, K. et al. Divergent clonal evolution of blastic plasmacytoid dendritic cell neoplasm and chronic myelomonocytic leukemia from a shared TET2-mutated origin. *Leukemia* **35**, 3299–3303 (2021).

17. Togami, K. et al. Sex-biased mutations in myeloid malignancies impair plasmacytoid dendritic cell activation and apoptosis. *Cancer Discov.* **12**, 522–541 (2022).

18. Menezes, J. et al. Exome sequencing reveals novel and recurrent mutations with clinical impact in blastic plasmacytoid dendritic cell neoplasm. *Leukemia* **28**, 823–829 (2014).

19. Patel, B. J. et al. Genomic determinants of chronic myelomonocytic leukemia. *Leukemia* **31**, 2815–2823 (2017).

20. Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).

21. Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).

22. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).

23. Xu, W. et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).

24. Figueroa, M. E. et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553–567 (2010).

25. Losman, J.-A. et al. (*R*)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science* **339**, 1621–1625 (2013).

26. van Galen, P. et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell* **176**, 1265–1281 (2019).

27. Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).

28. Sapienza, M. R. et al. Blastic plasmacytoid dendritic cell neoplasm: genomics mark epigenetic dysregulation as a primary therapeutic target. *Haematologica* **104**, 729–737 (2019).

29. Brash, D. E. UV signature mutations. *Photochem. Photobiol.* **91**, 15–26 (2015).

30. Strick, T. R. & Savery, N. J. Understanding bias in DNA repair. *Proc. Natl Acad. Sci. USA* **114**, 2791–2793 (2017).

31. Hanawalt, P. C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat. Rev. Mol. Cell Biol.* **9**, 958–970

(2008).

32. Hodis, E. et al. A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).

33. Redecke, V. et al. Hematopoietic progenitor cell lines with myeloid and lymphoid potential. *Nat. Methods* **10**, 795–803 (2013).

34. Dunkern, T. R., Fritz, G. & Kaina, B. Ultraviolet light-induced DNA damage triggers apoptosis in nucleotide excision repair-deficient cells via Bcl-2 decline and caspase-3/-8 activation. *Oncogene* **20**, 6026–6038 (2001).

35. Ostrander, E. L. et al. Divergent effects of *Dnmt3a* and *Tet2* mutations on hematopoietic progenitor cell fitness. *Stem Cell Rep.* **14**, 551–560 (2020).

36. Yin, Q. et al. Ultraviolet B irradiation induces skin accumulation of plasmacytoid dendritic cells: a possible role for chemerin. *Autoimmunity* **47**, 185–192 (2014).

37. Nestle, F. O. et al. Plasmacytoid predendritic cells initiate psoriasis through interferon-α production. *J. Exp. Med.* **202**, 135–143 (2005).

38. Gregorio, J. et al. Plasmacytoid dendritic cells sense skin injury and promote wound healing through type I interferons. *J. Exp. Med.* **207**, 2921–2930 (2010).

39. Heier, I. et al. Sun exposure rapidly reduces plasmacytoid dendritic cells and inflammatory dermal dendritic cells in psoriatic skin. *Br. J. Dermatol.* **165**, 792–801 (2011).

40. Machado, H. E. et al. Diverse mutational landscapes in human lymphocytes. *Nature* **608**, 724–732 (2022).

41. Brady, S. W. et al. The genomic landscape of pediatric acute lymphoblastic leukemia. *Nat. Genet.* **54**, 1376–1389 (2022).

42. Wang, L. et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet.* **47**, 1426–1434 (2015).

43. Lee, W. Y., Weinberg, O. K. & Pinkus, G. S. GATA1 is a sensitive and specific nuclear marker for erythroid and megakaryocytic lineages. *Am. J. Clin. Pathol.* **147**, 420–426 (2017).

44. Shanmugam, V. et al. Cyclin D1 is expressed in neoplastic cells of Langerhans cell histiocytosis but not reactive Langerhans cell proliferations. *Am. J. Surg. Pathol.* **41**, 1390–1396 (2017).

45. Kluk, M. J. et al. Validation and Implementation of a custom next-generation sequencing clinical assay for hematologic malignancies. *J. Mol. Diagn.* **18**, 507–515 (2016).

46. Garcia, E. P. et al. Validation of OncoPanel: a targeted next-generation sequencing assay for the detection of somatic variants in cancer. *Arch. Pathol. Lab. Med.* **141**, 751–758 (2017).

47. Sholl, L. M. et al. Institutional implementation of clinical tumor profiling on an unselected cancer population. *JCI Insight* **1**, e87062 (2016).

48. MacConaill, L. E. et al. Prospective enterprise-level molecular genotyping of a cohort of cancer patients. *J. Mol. Diagn.* **16**, 660–672 (2014).

49. Liu, D. et al. C/EBPβ enhances platinum resistance of ovarian cancer cells by reprogramming H3K79 methylation. *Nat. Commun.* **9**, 1739 (2018).

50. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

51. Broad Institute. Picard, v.2.5.0, http://broadinstitute.github.io/picard/ (2016).

52. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11.10.1–11.10.33 (2013).

53. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).

54. Alexandrov, L. B. et al. Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015).

55. Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med.* **10**, 33 (2018).

56. Hughes, T. K. et al. Second-strand synthesis-based massively parallel scRNA-seq reveals cellular states and molecular features of human inflammatory skin pathologies. *Immunity* **53**, 878–894 (2020).

57. DePasquale, E. A. K. et al. Single-cell multiomics reveals clonal T-cell expansions and exhaustion in blastic plasmacytoid dendritic cell neoplasm. *Front. Immunol.* **13**, 809414 (2022).

58. Miller, T. E. et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat. Biotechnol.* **40**, 1030–1034 (2022).

59. Nam, A. S. et al. Somatic mutations and cell identity linked by Genotyping of Transcriptomes. *Nature* **571**, 355–360 (2019).

60. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

61. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

62. Pullin, J. M. & McCarthy, D. J. A comparison of marker gene selection methods for single-cell RNA sequencing data. Preprint at *bioRxiv* https://doi.org/10.1101/2022.05.09.490241 (2022).

63. Wickham, H. et al. Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

Acknowledgements

We thank all of the patients and families for participation in this research study; A. Kreso and A. S. Nam for discussions; and P. Rogers for flow cytometry support. G.K.G. is supported by the Damon Runyon Cancer Research Foundation; V.H. by the Charles H. Hood Foundation, the Children's Cancer Research Fund and the V Foundation; P.v.G. by the National Cancer Institute (R00 CA218832), the Starr Cancer Consortium, the Gilead Sciences Research Scholars Program in Hem/Onc, the Harvard Medical School Epigenetics & Gene Dynamics Initiative, the Leukaemia & Lymphoma Society, the Glenn Foundation for Medical Research, the American Federation for Aging Research (AFAR) and the William Guy Forbeck Research Foundation; P.v.G. and A.A.L. by the Ludwig Center at Harvard and the Bertarelli Rare Cancers Fund; and A.A.L. by the NCI (R37 CA225191), Department of Defense (W81XWH-20-1-0683) and the Mark Foundation for Cancer Research. A.A.L is a Scholar of the Leukaemia & Lymphoma Society. None of the funding bodies were directly involved in the design of the study, nor in the collection, analysis or interpretation of the data, or writing of the manuscript. Some of the figure panels (Figs. 1a, 2c, 3a, 5b & 5f) were generated using BioRender.

Author information

Authors and Affiliations

1. Department of Pathology, Dana-Farber Cancer Institute, Boston, MA, USA

   Gabriel K. Griffin

2. Broad Institute of MIT and Harvard, Cambridge, MA, USA

   Gabriel K. Griffin, Daniel Ssozi, Julia A. Verga, Yoke Seng Lee, Vignesh Shanmugam, Bradley E. Bernstein, Volker

Hovestadt, Peter van Galen & Andrew A. Lane

3. Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

Gabriel K. Griffin, Vignesh Shanmugam, Jason L. Hornick & Elizabeth A. Morgan

4. Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

Christopher A. G. Booth, Katsuhiro Togami, Sun Sook Chung, Juliette M. Bouyssou & Andrew A. Lane

5. Division of Hematology, Brigham and Women's Hospital, Boston, MA, USA

Daniel Ssozi, Yoke Seng Lee & Peter van Galen

6. Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

Julia A. Verga & Bradley E. Bernstein

7. Department of Dermatology, Center for Cutaneous Oncology, Dana-Farber Cancer Institute and Brigham and Women's Hospital, Boston, MA, USA

Nicole R. LeBoeuf

8. Department of Cell Biology, Harvard Medical School, Boston, MA, USA

Bradley E. Bernstein

9. Ludwig Center at Harvard, Harvard Medical School, Boston, MA, USA

Bradley E. Bernstein, Peter van Galen & Andrew A. Lane

10. Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

    Volker Hovestadt

11. Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA

    Volker Hovestadt

Contributions

G.K.G., C.A.G.B., K.T., S.S.C., D.S., J.A.V., J.M.B., Y.S.L., V.S., J.L.H., N.R.L., E.A.M., B.E.B., V.H., P.v.G. and A.A.L. conducted experiments and analysed data. G.K.G., K.T., S.S.C., D.S., J.A.V., Y.S.L., B.E.B., V.H., P.v.G. and A.A.L. conducted sequencing experiments. G.K.G., V.S., J.L.H., N.R.L. and E.A.M. conducted clinical and pathological analyses. G.K.G., C.A.G.B., K.T., J.M.B. and A.A.L. generated mouse engineered haematopoietic cells and conducted in vitro differentiation experiments. G.K.G., V.H., P.v.G. and A.A.L. designed the study, interpreted the data and wrote the paper. All of the authors participated in editing the paper.

Corresponding authors

Correspondence to Gabriel K. Griffin, Volker Hovestadt, Peter van Galen or Andrew A. Lane.

Ethics declarations

Competing interests

B.E.B. discloses financial interests in Fulcrum Therapeutics, Chroma Medicine, HiFiBio, Arsenal Biosciences, Cell Signaling Technologies and Design Pharmaceuticals. A.A.L. received research funding from Abbvie and Stemline therapeutics, and consulting fees from Cimeio Therapeutics, IDRx, Jnana Therapeutics, N-of-One and Qiagen. P.v.G. received consulting fees from ManaT Bio and Immunitas.

Peer review

Peer review information

Nature thanks Eric Deconinck, George Vassiliou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Clinical evaluation of bone marrow involvement in BPDCN.

a, Clinical photo of a representative BPDCN skin lesion at initial diagnosis showing a violaceous nodule that occurred on a background of diffuse pink, brown and violaceous papules, plaques, and tumours. **b**, Low-power histologic images of a representative BPDCN skin tumour showing extensive dermal infiltration by malignant BPDCN cells with characteristic immunophenotype, including positivity for TCL1, CD4, CD56, and CD123, and negativity for CD34, myeloperoxidase (MPO), and lysozyme (LYZ). These features distinguish BPDCN in the skin from histologic mimics, including acute myeloid leukaemia with monocytic differentiation and chronic myelomonocytic leukaemia. **c**, Immunophenotypic features of the sixteen BPDCN cases included in the cohort, as determined by immunohistochemistry and/or flow cytometry performed during initial pathology evaluation. This shows the presence of BPDCN markers (TCL1, CD4, CD56, CD123) and the absence of myeloid, T, or B-cell lineage markers that define other forms of acute leukaemia. **d**, Tumour phylogenies and sample-specific alterations for Patients 1 and 12 as determined by whole-genome sequencing. Somatic SNVs (red) and copy-number alterations (blue) were defined at the latest relapse time point, and then

assessed in prior skin tumour and bone marrow diagnosis samples. **e**, Tumour phylogenies and sample-specific alterations for Patients 2, 3, 5, and 8 as determined by targeted sequencing (RHP and Oncopanel, see Methods). Mutations that are not covered in both the RHP (bone marrow samples) and Oncopanel (skin samples) assays are indicated in parentheses. **f**, Immunohistochemistry for MTAP protein (brown) as a marker of chr 9p21.3 loss in BPDCN skin tumours. MTAP is lost in skin tumour cells from patient 2, which contain a focal deletion on chr9p21.3, and retained in skin tumour cells from patient 3, which do not have chr9p21.3 loss. **g**, Inferred clonal architecture in diagnostic marrows for Patients 1 and 12 in (d). Premalignant subclones related to BPCN skin tumours are indicated with arrows, and further annotated with skin-specific progression mutations. **h**, Frequency of copy-loss in CDKN2A/chr9p and SETD2/chr3p in BPDCN and AML. Statistical significance by two-sided Fisher's exact test. TCGA, The Cancer Genome Atlas. DFCI, Dana-Farber Cancer Institute. ** $P < 0.01$, **** $P < 0.0001$. Related to Fig. 1.

Extended Data Fig. 2 Copy-number alteration analysis of Patient 7 and 9 whole-exome sequencing data.

a, Genome plots show B (minor) allele frequencies (left) and read coverage (right) for single-nucleotide variants (SNVs) detected in the Patient 7 germline sample using whole-exome sequencing. Heterozygous SNVs (A/B) are indicated in red, homozygous SNVs (B/B) are indicated in green. The same SNVs are shown for the bone marrow and skin tumour samples, indicating copy-number alterations specific to the skin tumour. A focal loss detected on chr2q is indicated by arrows. **b**, Genome plots show similar analysis for germline, bone marrow, and skin tumour samples of Patient 9. A sub-clonal copy-number neutral loss of heterozygosity (LOH) on chr7q was detected in the uninvolved bone marrow sample, which was not detected in the skin tumour. Related to Fig. 1.

Extended Data Fig. 3 Copy-number alteration analysis of Patient 10 whole-exome sequencing data.

a, Genome plots show B (minor) allele frequencies (left) and read coverage (right) for single-nucleotide variants (SNVs) detected in the Patient 10

germline sample using whole-exome sequencing. Heterozygous SNVs (A/B) are indicated in red, homozygous SNVs (B/B) are indicated in green. The same SNVs are shown for bone marrow and skin tumour sample #1 and #2, indicating copy-number alterations specific to the skin tumours. A homozygous deletion of the *CDKN2A* gene locus is indicated by arrows. An additional focal loss on chr4q in skin tumour #2 is also indicated. **b**, Genome plots show SNVs detected in the relapse skin tumour and bone marrow samples of Patient 10 collected after receiving an allogeneic stem cell transplant. SNVs homozygous in the host germline sample are indicated in blue (A/A) and green (B/B). Heterozygous SNVs are indicated in red (A/B). A low level of donor DNA is detected in the relapse skin tumour sample (<5%). The estimated percentage of donor DNA is 60% in the relapse bone marrow sample. Copy-number alterations specific to the skin tumours (i.e. loss of chr9, chr3p, and copy-number neutral LOH of chr6p) are detected in both relapse samples. The only additional alteration is a focal loss on chr1q. Related to Fig. 1.

Extended Data Fig. 4 Copy-number alteration analysis of Patient 1, 3, and 12 whole-genome sequencing data.

a, Genome plots show B (minor) allele frequencies for all single-nucleotide variants (SNVs) detected in the Patient 1 germline sample using whole-genome sequencing. SNVs are shown for bone marrow samples at diagnosis and remission, and a skin tumour sample. SNVs that are heterozygous (A/B) in the germline sample are indicated in red, homozygous SNVs (B/B) are indicated in green. Loss of chr9 specific to the skin tumour sample is indicated. **b**, Genome plot shows similar analysis for a bone marrow sample of Patient 3, not harbouring any copy-number alterations. **c**, Genome plots show similar analysis for a bone marrow and skin tumour sample of Patient 12. Copy-number neutral loss of heterozygosity (LOH) on chr4q was detected in both samples. Additional loss of chr15q and gain of chr1 was detected in the skin tumour sample. **d**, Genome plots show single-nucleotide variants (SNVs) detected in the relapse bone marrow sample of Patient 1 collected after receiving an allogeneic stem cell transplant. SNVs are shown for skin relapse sample #1 and #2, as well as a bone marrow relapse. Samples were collected at consecutive time points. SNVs homozygous in the host germline sample are indicated in blue (A/A) and green (B/B),

heterozygous SNVs are indicated in red (A/B). In addition to the loss of chr9 in the skin sample at diagnosis, additional copy-number alterations are detected in each sample: Loss of chr3p is first detected in skin relapse #1, loss of chr2p, chr4q, and chr17p are first detected in skin relapse #2, and loss of chr5q is first detected in the bone marrow relapse. Donor DNA is detected in both skin sample #2 and bone marrow relapse sample. Estimated fraction of donor DNA is <5% and 40%, respectively. **e**, Genome plot shows similar analysis for a bone marrow relapse sample of Patient 12, who did not receive an allogeneic stem-cell transplant. Additional alterations in chr1, chr5q, and chr12p are indicated. Related to Fig. 1.

Extended Data Fig. 5 Single-cell transcriptome analysis of healthy donor and BPDCN patient bone marrow samples.

a, Heatmap depicts five-fold cross validation of the random forest classifier comprising 22 classes corresponding to the cell types identified in healthy bone marrow samples (including inferred cell doublets). The healthy donor cells were split such that 80% were used as a reference to predict the classification of the remaining 20%. This process was repeated five times. Cells that fall on the diagonal are classified according to their annotation. Cells that do not fall on the diagonal are mis-classified as a different cell type. **b**, Heatmaps show expression of known marker genes (rows) in cells (columns) of the myeloid and erythroid differentiation trajectories in bone marrow samples from healthy donors (left) and from BPDCN patients without known bone marrow involvement (right). Top annotation bars show cell types (colour legend provided in panel c) and sample identity. **c**, Barplot shows the proportion of cell types within each of the 17 analysed single-cell samples. Healthy donor cells were annotated by clustering and assessment of marker gene expression. Patient cells were annotated using the random forest classifier, using healthy donors as a reference. The high proportion of T cells in Patient 10 was consistent with flow cytometry. **d**, Dot plot shows percent of cells classified as pDCs in each healthy donor and patient sample. High percentage of pDCs in samples with known marrow involvement reflects malignant BPDCN cells. Data is shown for all samples that were analysed by scRNA-seq (n = 6 healthy donors, n = 5 samples without known marrow involvement, and n = 6 samples with marrow involvement). Statistical significance between sample groups is indicated (two-sided

Student's t-test). **e**, Scatterplot shows B (minor) allele frequencies of Patient 10 SNVs in relapse bone marrow sample (post allogeneic stem cell transplant, y-axis) and germline sample (x-axis). SNVs homozygous in the host germline sample are indicated in blue (A/A) and green (B/B), heterozygous SNVs are indicated in red (A/B). This analysis allows for the identification of alleles that are specific for host and donor cells (indicated in bold). Thresholds that were used for subsetting these SNVs are indicated. **f**, Scatterplot shows quantification of host- and donor-specific alleles in the scRNA-seq data of Patient 10 relapse bone marrow samples (red). The fraction of SNVs specific to the donor genome is indicated on the x-axis. Genotypes could be assigned for the majority of cells (98.0%), with 62.6% of those annotated as host-derived, and 37.4% annotated as donor-derived. Cells from the diagnostic bone marrow sample (blue), for which no cells were annotated as donor-derived, are shown as comparison. **g**, Barplot shows the proportion of cell types within Patient 10 relapse cells genotyped as host- and donor-derived. Most host cells classify as pDCs, likely reflecting malignant BPDCN cells. **h**, Violin plot shows scores of a published BPDCN gene signature for pDCs from each of the 17 analysed single-cell samples. Related to Fig. 2.

Extended Data Fig. 6 Single-cell genotyping of bone marrow samples.

a, Scatterplot shows overview of the genotyping efficiency of all 40 mutations targeted by XV-seq across 11 samples. Some mutations were analysed in multiple samples collected from the same patient. **b**, Genome plot shows combined 10x scRNA-seq reads for Patient 10 relapse bone marrow donor cells and host cells over the *CDKN2A* gene locus on chromosome 9. The focal homozygous deletion observed in (malignant) host cells results in atypical splicing of the upstream *MTAP* gene to five different acceptor sites downstream of *CDKN2A*. This enabled the generation of a genotyping primer specific to exon 4 of *MTAP* that can be used to detect the *CDKN2A* deletion in single cells. **c**, Scatterplot compares the number of genotyped cells detected in raw scRNA-seq data (x-axis) with the number of genotyped cells detected by XV-seq (y-axis, $r = 0.71$). Median enrichment across targets is 11.1-fold (indicated by dashed line). These data demonstrate XV-seq target enrichment and the utility of selecting suitable mutations

based on raw scRNA-seq data. **d**, Scatterplot shows the genotyping efficiency of XV-seq targets (y-axis) compared to the normalized expression level of the transcript (x-axis, r = 0.55). **e**, Scatterplot shows agreement between VAFs from bulk targeted sequencing using the Rapid Haem Panel (y-axis) and single-cell genotyping using XV-seq (x-axis, r = 0.73). For the latter, the VAF was calculated as the number of mutated transcripts / number of total transcripts captured. **f**, Barplot shows the percentage of cells from the Patient 10 relapse sample (post stem cell transplant) for which genotype information was obtained (Extended Data Fig. 5e–g). *RAB9A* is located on chromosome X (male patient) and *CDKN2A* is located on chromosome 9 of which one copy is lost in addition to the focal deletion of the locus (Extended Data Fig. 3a). The exclusive detection of wild-type and mutated transcripts in the expected cell populations supports accuracy of cell type annotation, host/donor classification, and XV-seq mutation detection. **g**, Illustration of supporting evidence for subclonal structure obtained from single-cell XV-seq of the Patient 10 uninvolved bone marrow. *TET2* mutations S792* and Q1034* co-occur in the same cell (major subclone). Similarly, *TET2* mutations H1216* and H1380Y also co-occur in the same cell (minor subclone). Mutations specific to the two subclones were not detected in the same cell. The sample is karyotypically normal, further supporting the existence of two subclones, as truncating mutations in *TET2* are unlikely to affect the same allele. VAFs from targeted sequencing are indicated between parentheses. **h**, Illustration of supporting evidence for subclonal structure obtained from single-cell XV-seq of the Patient 9 uninvolved bone marrow. This sample is characterized by a sub-clonal loss of heterozygosity (LOH) on chr7q. Detection of the lost haplotype in n=25 single cells indicates that mutations in *TET2* and *CUX1* occurred before the LOH of chr7q. VAFs from targeted sequencing are indicated between parentheses. The high VAF of the mutation in *CUX1* is explained by its location on chr7q. **i**, Heatmaps show proportion of *TET2*-mutated cells in each major hematopoietic cluster. Ten patient-specific *TET2* mutations were assessed in five marrow samples. *P*-values indicate mutant-cell enrichment in HSPC/Erythroid/Myeloid vs. B/T/NK cells by Pearson's Chi-square test with Yates' correction. Related to Fig. 2.

Extended Data Fig. 7 Single cell-derived BPDCN signature enables malignant cell classification.

a, UMAPs show all cells classified as pDCs across the entire scRNA-seq dataset (n = 14,430 cells), coloured according to patient/sample (left), samples with and without known BPDCN involvement (middle), and BPDCN signature score (right). **b**, Volcano plot shows differentially expressed between healthy pDCs and malignant BPDCN cells. We used 45 genes with log2 fold change > 1 and adjusted *P*-value < 1E-30 (green symbols), including *BCL2* and *TCL1A*, to calculate the BPDCN signature score in downstream analyses. All 45 genes are provided in Supplementary Table 3c. *P*-values were calculated using a two-sided Wilcoxon Rank Sum test and adjusted using Bonferroni correction as implemented in the Seurat function FindMarkers. **c**, Scatterplot shows signature scores in all cells that were classified as pDC (n = 14,430). We calculated scores using a previously published BPDCN module[27] (x-axis) and using the 45 signature genes we defined in panel a (y-axis). **d**, Sinaplot shows BPDCN signature scores in all single cells (n = 87,011) that we profiled in this study, split by donor type. The colour indicates cell type classification by the RandomForest algorithm. **e**, Scatterplots show single cells from all patient samples (n = 66,600) according to their random forest pDC prediction score (x-axis) and BPDCN signature score (y-axis). Red dots indicate detection of mutant transcripts (n = 16 founder mutations, left; n = 15 progression mutations, right), grey dots indicate detection of wild-type transcripts. **f**, Sinaplot shows BPDCN signature scores in all single cells (n = 87,011), split by cell type and coloured by final classification. **g**, Violin plots show expression of canonical marker genes in cells that were originally classified as proB cells and plasma cells. Cells with a BPDCN signature score exceeding 0.5 were reclassified as malignant cells. The absence of CD19 in reclassified proB cells and the absence of CD138 in reclassified plasma cells supports our reclassification. **h**, Flow chart illustrates classification of healthy pDCs, premalignant pDCs, and malignant BPDCN cells for single cells across the dataset. Related to Fig. 2.

Extended Data Fig. 8 Gene expression and mutation analysis outlines BPDCN disease progression.

a, Scatterplot shows gene expression fold changes between premalignant vs. healthy pDCs (x-axis) and malignant BPDCN cells vs. healthy pDCs (y-axis). The changes are positively correlated (r = 0.28) and more pronounced

for the malignant BPDCN cell comparison. **b**, Violin/sina plots show expression of selected genes that are differentially expressed between healthy pDCs (n = 203), premalignant pDCs (n = 495), and malignant BPDCN cells (n = 14,232). Cells are grouped according to cell annotations defined in Extended Data Fig. 7f. Adjusted *P*-values indicate a comparison to healthy pDCs and were calculated using a two-sided Wilcoxon Rank Sum test and adjusted using Bonferroni correction as implemented in the Seurat function FindAllMarkers. **c**, Sina plots show the BPDCN signature score in all cells classified as pDC in bone marrow samples from healthy donors and patients without involvement. Red dots indicate cells with detection of mutant transcripts, green dots indicate cells with detection of wild-type transcripts (n = 16 founder mutations, left; n = 13 progression mutations, right). Progression mutations were nearly exclusively captured in cells with a BPDCN signature score exceeding 0.5, consistent with the malignant identity of these cells (*P* = 1.1E-84 by Pearson's Chi-square test with Yates' correction). **d**, Scatterplots show single cells from the Patient 10 diagnostic sample (n = 10,106) according to their random forest pDC prediction score (x-axis) and BPDCN signature score (y-axis). Red dots indicate detection of mutant transcripts, green dots indicate detection of wild-type transcripts for six indicated genes. Integrated analysis of gene expression and mutations concertedly identified rare circulating malignant BPDCN cells (n = 19, 0.19%). **e**, Scatterplot shows single cells from the Patient 12 diagnostic sample (n = 6,862) according to their random forest pDC prediction score (x-axis) and BPDCN signature score (y-axis). Red dots indicate detection of mutant transcripts, green dots indicate detection of wild-type transcripts for *MALAT1*. The detection of mutated transcripts in two cells with a high BPDCN signature score supports their malignant identity. Related to Fig. 2.

Extended Data Fig. 9 Mutational signature analysis of BPDCN patient samples.

a, Heatmaps show mutational signature analysis of BPDCN patient samples analysed by whole-exome sequencing from three published datasets[16,17,28]. Blue heat indicates the predicted relative contribution of the mutational signature. Relative contribution of signature 7 (UV damage signature) is indicated. Samples from Batta et al. were generated from indicated bone marrow populations at diagnosis (n = 3) and relapse (n = 1), and a skin

tumour sample, all from the same patient. The UV damage signature is detected in all samples containing malignant cells (relative contribution ≥0.38). Samples from Sapienza et al. were generated from skin tumours of nine patients. The UV damage signature is detected in five samples (≥0.27). Samples from Togami et al. were generated from sorted malignant cells of bone marrows from 11 patients. The UV damage signature is detected in four samples (≥0.20). **b**, Heatmaps show sequence context of somatic SNVs detected in samples analysed by whole-exome sequencing. Mutations are only shown in the sample in which they were first detected, and not in subsequent samples. Column headers indicate the base substitution. X- and y-axis labels indicate the bases upstream and downstream of the mutated base, respectively. Red heat indicates the relative contribution of mutations within the given nucleotide context. UV-associated CC > CT and TC > TT mutations are indicated by bold horizontal boxes and their total percentages. **c**, Heatmaps show sequence context of somatic SNVs detected in samples analysed by whole-genome sequencing. Annotations are similar to panel b. **d**, Barplot shows the number of dinucleotide mutations detected in samples from Patient 1, grouped by UV-specific CC > TT and all other dinucleotide mutations. Mutations are only shown in the sample in which they were first detected. **e**, Barplots indicate the percentage of cells for which wild-type or mutant transcript were detected using XV-seq. Selected UV-associated (CC > CT and TC > TT) as well as UV-specific (CC > TT, ETV6) progression mutations are shown. Cells are grouped according to cell annotations defined in Extended Data Fig. 7f. Notably, in the diagnostic sample of Patient 10 (without known marrow involvement), UV-associated mutations are specifically detected in rare malignant cells (n = 19), consistent with our model of retrograde dissemination. Related to Fig. 3.

Extended Data Fig. 10 Order of acquisition of UV damage and functional evaluation in the HOXB8 model of pDC differentiation.

a, Scatterplot shows B (minor) allele frequencies of SNVs located on chromosome 3 in skin tumour #1 and #2 from Patient 10. SNVs are coloured according to their genotype in the matching germline sample. Loss of a region on chr3p (Extended Data Fig. 3a) affects different alleles, indicating separate events and convergent evolution in both skin tumours samples.

Heterozygous SNVs were phased along the diagonal (shown in Fig. 4d). **b**, Barplots show VAFs for founder and progression mutations detected in bone marrow samples (top) and matched skin tumour samples (bottom) from Patients 7 ans 9. UV-associated TC > TT mutations are indicated in orange, other mutations in green. Average VAF (solid line) and coefficient of variation (CV, standard deviation divided by mean, dotted lines) of founder mutations are indicated. Asterisks indicate mutations that are affected by copy-number alterations (Patient 9) or are likely contaminants from blood (Patient 7) and are not included in this calculation. **c**, UMAPs of scRNA-seq performed on bone marrow cells (n=7,374 cells) from Patient 14. Cell type clusters (left), BPDCN signature scores (middle), and single cell genotyping results for the UV-specific (CC>TT) ETV6.R369W mutation (right) are shown. **d**, Flow cytometry gating strategy for the identification of pDCs (CD11c+/CD11b-/B220+) and cDCs (CD11c+/CD11b+/B220-) from representative HOXB8 cultures. **e**, Time course showing the dynamics of representative pDC and cDC differentiation from day 3 to day 8 following oestrogen withdrawal in the HOXB8 system. pDC and cDC populations become distinguishable on days 6-8. **f**, Amplicon sequencing showing successful CRISPR/Cas9 editing of *Tet2* in HOXB8 cells using two different gRNAs. Greater than 95% of reads for each guide are edited, including insertions (red boxes) and deletions (dashes) predicted to cause frameshift alterations. The predicted Cas9 cleavage site 3 nucleotides upstream from the NGG PAM site is indicated with a vertical dashed line. Related to Fig. 4.

Supplementary information

Reporting Summary

Peer Review File

Supplementary Table 1

Overview of patient characteristics and datasets. Demographic and clinical information for all of the patients (*n* = 16) included in the study (Supplementary Table 1a) and an overview of the genomic assays performed (Supplementary Table 1b).

Supplementary Table 2

Genome sequencing results. Summary of targeted (Supplementary Table 2a), whole-exome (Supplementary Table 2b–d) and whole-genome (Supplementary Table 2e) sequencing results for all patients ($n = 16$, $n = 3$ and $n = 3$, respectively) included in the study. Multiple samples and/or timepoints were analysed for most patents (total of $n = 32$, $n = 12$ and $n = 13$, respectively, including germline samples). Tables for whole-exome sequencing results, providing genomic location, sequence context, VAF and gene annotations for all high-confidence mutations.

Supplementary Table 3

scRNA-seq analysis. Clinical information and quality metrics for scRNA-seq analysis of 6 healthy donors and 11 patient samples (Supplementary Table 3a). The top 50 marker genes for each of the 21 cell types identified by clustering of healthy bone marrow samples (Supplementary Table 3b). 45 genes that are significantly upregulated in malignant BPDCN cells compared with in healthy pDCs (Supplementary Table 3c).

Supplementary Table 4

Single-cell genotyping primers. Oligonucleotides that were used for XV-seq, including site-specific primers (Supplementary Table 4a), universal primers for Seq-Well S^3 (Supplementary Table 4b) and universal primers for 10x v.3 3′ libraries (Supplementary Table 4c).

Source data

Source Data Fig. 1

Source Data Fig. 5

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Griffin, G.K., Booth, C.A.G., Togami, K. *et al.* Ultraviolet radiation shapes dendritic cell leukaemia transformation in the skin. *Nature* **618**, 834–841 (2023). https://doi.org/10.1038/s41586-023-06156-8

- Received: 21 January 2021
- Accepted: 02 May 2023
- Published: 07 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06156-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article is cited by

- **[Ultraviolet light shapes the evolution of precancerous cells]**
    - Elli Papaemmanuil

  Nature (2023)

Ultraviolet light shapes the evolution of precancerous cells

- Elli Papaemmanuil

Nature News & Views 07 Jun 2023

Enlightening a leukemia

- Nicholas J. Bernard

Nature Immunology Research Highlight 20 Jun 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06156-8

- Article
- Open Access
- Published: 31 May 2023

Engineered tRNAs suppress nonsense mutations in cells and in vivo

- Suki Albers,
- Elizabeth C. Allen,
- Nikhil Bharti,
- Marcos Davyt,
- Disha Joshi,
- Carlos G. Perez-Garcia,
- Leonardo Santos,
- Rajesh Mukthavaram,
- Miguel Angel Delgado-Toscano,
- Brandon Molina,
- Kristen Kuakini,
- Maher Alayyoubi,
- Kyoung-Joo Jenny Park,
- Grishma Acharya,
- Jose A. Gonzalez,
- Amit Sagi,
- Susan E. Birket,
- Guillermo J. Tearney,
- Steven M. Rowe,
- Candela Manfredi,
- Jeong S. Hong,
- Kiyoshi Tachikawa,
- Priya Karmali,
- Daiki Matsuda,

- Eric J. Sorscher ,
- Pad Chivukula &
- …
- Zoya Ignatova

Show authors

Nature **volume 618**, pages 842–848 (2023)

- 15k Accesses
- 178 Altmetric
- Metrics details

Abstract

Nonsense mutations are the underlying cause of approximately 11% of all inherited genetic diseases[1]. Nonsense mutations convert a sense codon that is decoded by tRNA into a premature termination codon (PTC), resulting in an abrupt termination of translation. One strategy to suppress nonsense mutations is to use natural tRNAs with altered anticodons to base-pair to the newly emerged PTC and promote translation[2,3,4,5,6,7]. However, tRNA-based gene therapy has not yielded an optimal combination of clinical efficacy and safety and there is presently no treatment for individuals with nonsense mutations. Here we introduce a strategy based on altering native tRNAs into efficient suppressor tRNAs (sup-tRNAs) by individually fine-tuning their sequence to the physico-chemical properties of the amino acid that they carry. Intravenous and intratracheal lipid nanoparticle (LNP) administration of sup-tRNA in mice restored the production of functional proteins with nonsense mutations. LNP–sup-tRNA formulations caused no discernible readthrough at endogenous native stop codons, as determined by ribosome profiling. At clinically important PTCs in the cystic fibrosis transmembrane conductance regulator gene (*CFTR*), the sup-tRNAs re-established expression and function in cell systems and patient-derived nasal epithelia and restored airway volume homeostasis. These results provide a framework

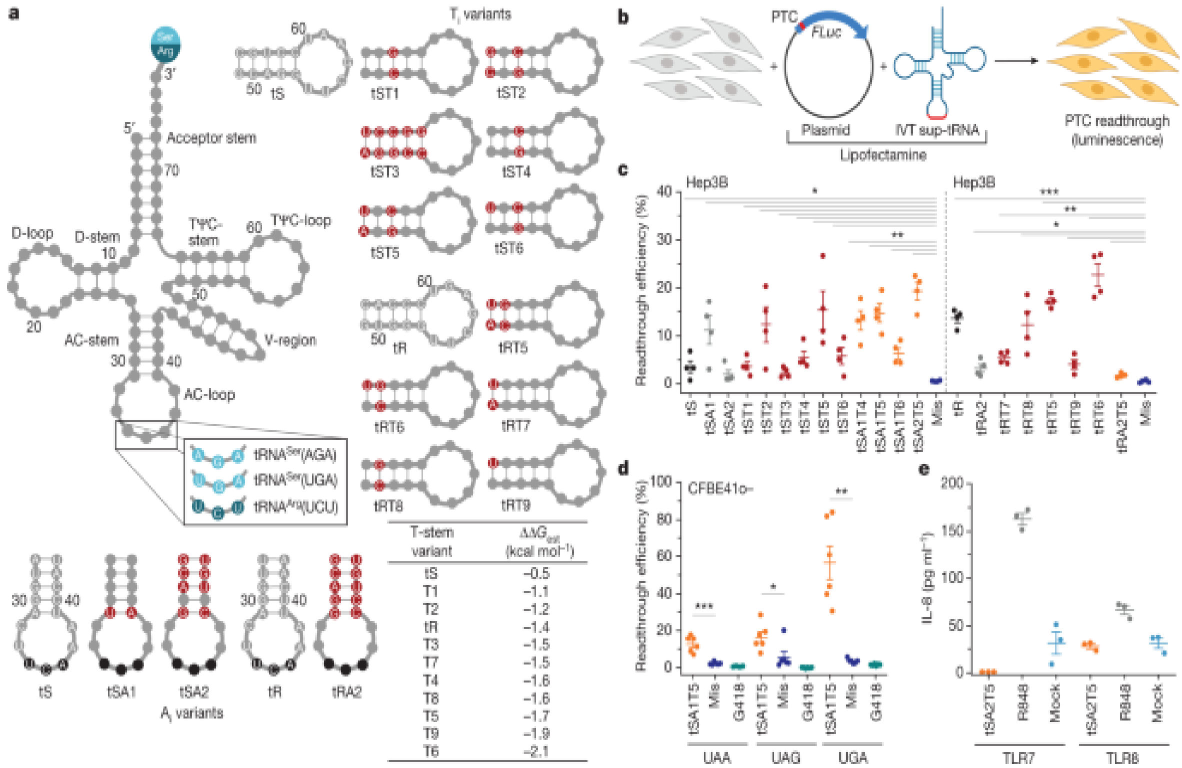
for the development of tRNA-based therapies with a high molecular safety profile and high efficacy in targeted PTC suppression.

Main

Efforts to develop treatments for patients with nonsense mutations focus on using low molecular weight pharmacological compounds or sup-tRNAs that induce readthrough at PTCs and restore translation and production of full-length proteins[2,3,4,5,6,7,8,9]. Although some pharmacological approaches have been used in clinical trials, non-specific insertion of random amino acids[10], off-target effects at natural stop codons and safety in long-term applications have limited their clinical use. Pioneered four decades ago[3,6], natural sense-codon-decoding tRNAs with an altered anticodon to decode stop codons have been shown to correct PTCs, but tRNA-based therapies based on this principle have not reached clinical trials because of insufficient efficacy, inability to achieve a therapeutic threshold and insufficient safety. One impediment is that not every native tRNA can be engineered into a sup-tRNA by altering its anticodon[11], in part because the decoding at a sense codon markedly differs from the hydrolysis at a stop codon[12], which is mediated by release factor[13] (eRF1) in eukaryotes. Moreover, newly emerged PTCs activate mRNA surveillance pathways (including nonsense-mediated mRNA decay (NMD)) to degrade mutated mRNA[14,15,16]. An anticodon-altered sup-tRNA may not establish the ideal geometry for decoding[11] to efficiently outcompete premature translation termination and the mRNA degradation process. Harnessing functionally conserved features of natural tRNAs and modulating sequences outside the anticodon, we successfully repurposed bacterial tRNAs to incorporate exclusively alanine at the UGA stop codon with codon–anticodon interactions resembling the Watson–Crick geometry of a sense-codon-decoding tRNA[17]. We reasoned that by applying a similar strategy and modulating various sequence segments of human tRNAs that are crucial for tRNA function in translation (Fig. 1a) such as the anticodon-stem and anticodon-loop—which modulate the accuracy of decoding—and the TΨC-stem that determines binding affinity to elongation factor[18,19,20] (eEF1A in humans), we would enhance sup-tRNA efficacy. We also leveraged a synthetic LNP system[21] to

encapsulate sup-tRNA and produce safe LNP–sup-tRNA with high efficacy for PTC suppression in vivo with a robust molecular safety profile.

Fig. 1: Sup-tRNA variants suppress different PTCs at Ser and Arg codons.



a, Schematic of a generic tRNA with the natural anticodon of human tRNA^{Ser}UGA, tRNA^{Ser}AGA or tRNA^{Arg}UCU (left). Nucleotide substitutions in the anticodon (AC)-stem (A_i variants; bottom) or the TΨC-stem (T_i variants; right) of tS or tR variants are highlighted in red. The table shows estimated ΔΔG values for binding affinities of eEF1A to the TΨC-stem. **b**, Schematic of the screening of sup-tRNAs with plasmid constructs encoding firefly luciferase (*FLuc*) with a PTC mutation (PTC-*FLuc*). IVT, in vitro-transcribed tRNA. **c**, Suppression efficacy of tS or tR variants in human liver Hep3B cells at FLuc(R208X) (in which X represents UGA), normalized to wild-type FLuc. Mis, mismatch tRNA. Data are mean ± s.e.m. ($n = 4$ independent replicates). **d**, tSA1T5 targets PTCs with different stop codon identities, as tested with Fluc(S466X) in human bronchial epithelial CFBE41o^{-} cells and normalized to wild-type FLuc. G418 (geneticin) is a low molecular mass readthrough-promoting agent. Data are mean ± s.e.m. ($n = 6$ independent replicates). **e**, TLR-dependent activation by tSA2T5 was

monitored in human TLR-transformed HEK293 cells. R848 agonist, which activates both TLR7 and TLR8, served as a positive control. Mock, mock-transfected cells. Data are mean ± s.e.m. ($n = 3$ independent replicates). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. One-sided t-test.

Source Data

Engineering sup-tRNAs for high efficacy

We selected three human $tRNA^{Ser}$, $tRNA^{Arg}$ and $tRNA^{Gly}$ families that decode codons that are frequently mutated to PTCs[1]. We first exchanged their anticodons to pair to the UGA PTC, resulting in tS, tR and tG, respectively (Fig. 1a, Extended Data Fig. 1 and Supplementary Table 1). For the initial screen, the in vitro-transcribed sup-tRNAs with functionally homogenous 3′ ends (Extended Data Fig. 2) were co-transfected in human Hep3B cells with a plasmid-encoded PTC reporter of firefly luciferase (*FLuc*) (Fig. 1b); the coding sequence of the *FLuc* reporter was extended at its 5′ end by 15 codons representing the sequence context of the most common PTC ($FLuc^{R208X}$; Supplementary Table 2) in the tripeptidyl peptidase 1 gene associated with lysosomal storage disorder. As expected, tS, tR and tG exhibited low readthrough activity, with tR showing the highest efficiency (Fig. 1c and Extended Data Fig. 1), probably because the natural $tRNA^{Arg}$UCU—the precursor of tR—is intrinsically prone to miscoding[22]. To achieve similar decoding efficiency among all natural tRNAs, their sequences have been fine-tuned by evolution to the chemical nature of the cognate amino acid, whereby the destabilizing thermodynamic effect of some amino acids is compensated by stronger interactions of the elongation factor with the TΨC-stem and vice versa[23]. The three selected tRNA families were aminoacylated with serine, arginine or glycine, which span the entire spectrum of thermodynamic contributions (that is, glycine is a destabilizing amino acid, serine is stabilizing and arginine is nearly neutral[23]). Next, we subjected tS, tR and tG to a comprehensive set of sequence changes, thereby preserving the recognition signals for the cognate aminoacyl-tRNA-synthetase (Fig 1a, Extended Data Fig. 1a and Supplementary Table 1). Changes in the TΨC-stem were made to stabilize (lower free energy difference ($\Delta\Delta G$) values) or destabilize (higher $\Delta\Delta G$

values) interactions with eEF1A (Fig. 1a, bottom right) and the energy contribution of different base pairs (positions 49–65, 50–64 and 51–63) were taken from ref. [23] (Methods). Position-specific changes in either anticodon-stem (A_i variants) or TΨC-stem (T_i variants) enhanced the readthrough efficiency of tS, and the simultaneous modulation of both anticodon-stem and TΨC-stem displayed the most robust effect (for example, variants tSA1T5 and tSA2T5; Fig 1c). The tS variants with a TΨC-stem interacting less stably with eEF1A (such as tST2 and tST5) exhibited higher suppression efficacy than tST6 (Fig. 1c), which had the most stable interaction with eEF1A (Fig. 1a, bottom right table). Enhancement of the suppression efficacy of sup-tRNA charged with serine (a stabilizing amino acid[23]) was achieved by modest stabilization of the interactions with eEF1A.

Arginine makes nearly no thermodynamic contribution to the aminoacyl-tRNA stability[23]. tR variants with substitutions in the TΨC-stem, which would stabilize the interactions with eEF1A, exhibited higher readthrough efficiency (that is, tRT6 the highest and tRT9 the lowest), whereas changes within the anticodon-stem alone (tRA2) or in combination with the TΨC-stem (tRA2T5) markedly reduced the PTC suppression (Fig. 1c). Glycine is a destabilizing amino acid[23], however mutations in the TΨC-stem stabilizing the interactions with eEF1A (tGT6) marginally enhanced the tG readthrough efficiency (Extended Data Fig. 1 and Supplementary Table 1). Overall, the tG variants were less effective than the tS and tR variants, probably owing to the intrinsic hyper-accuracy of natural glycine tRNAs[22]. Yet, for some PTCs this might be still sufficient to potentially address diseases in which the therapeutic threshold is low, such as cystic fibrosis[24]. Together, our findings indicate that to engineer native tRNAs to decode PTCs, unique design principles should be established for each tRNA family to trim accuracy in decoding and affinity for eEF1A tailored to the chemical nature of the cognate amino acid.

sup-tRNA efficacy at different PTCs

For one of the optimal tRNA variants with the highest readthrough activity (tSA1T5), which incorporated predominantly Ser at the UGA stop codon (Extended Data Fig. 3), we next assessed the suppression efficacy at UGA,

UAG and UAA as pathogenic nonsense mutations lead to all three PTC identities[1] (with a frequency of 38.5% for UGA, 40.4% for UAG and 21.1% for UAA). Within the screening process, to test the sup-tRNA efficacy in another pathogenic mutation context, we considered the S466X mutation in *CFTR*, which occurs naturally with different stop codon identities (https://cftr2.org/). Since this mutation is implicated in cystic fibrosis, we considered the human bronchial epithelial CFBE41o^{-} cell line—an established model in the cystic fibrosis drug expansion pipeline[25]. At an optimal sup-tRNA concentration, which does not alter cell viability (Extended Data Fig. 4a,b), tSA1T5 suppressed both UGA and UAG PTCs with higher efficiency than UAA PTCs (Fig. 1d and Extended Data Fig. 4c,d). Notably, tSA1T5 was substantially more efficient at all PTCs with three different stop codon identities than the known readthrough-stimulating antibiotic G418 (Fig. 1d). We also observed fourfold higher suppression at lower tSA1T5 doses when the sup-tRNA was co-administered with the PTC reporter as mRNA than as DNA (Extended Data Fig. 4e). The efficacy of tSA1T5 at the UGA PTC was greater for the S466X than for R208X (compare Fig. 1d with Fig. 1c), implying that the PTC sequence context also modulates sup-tRNA efficacy—an effect that has been reported for aminoglycosides-stimulated readthrough at natural termination codons[26].

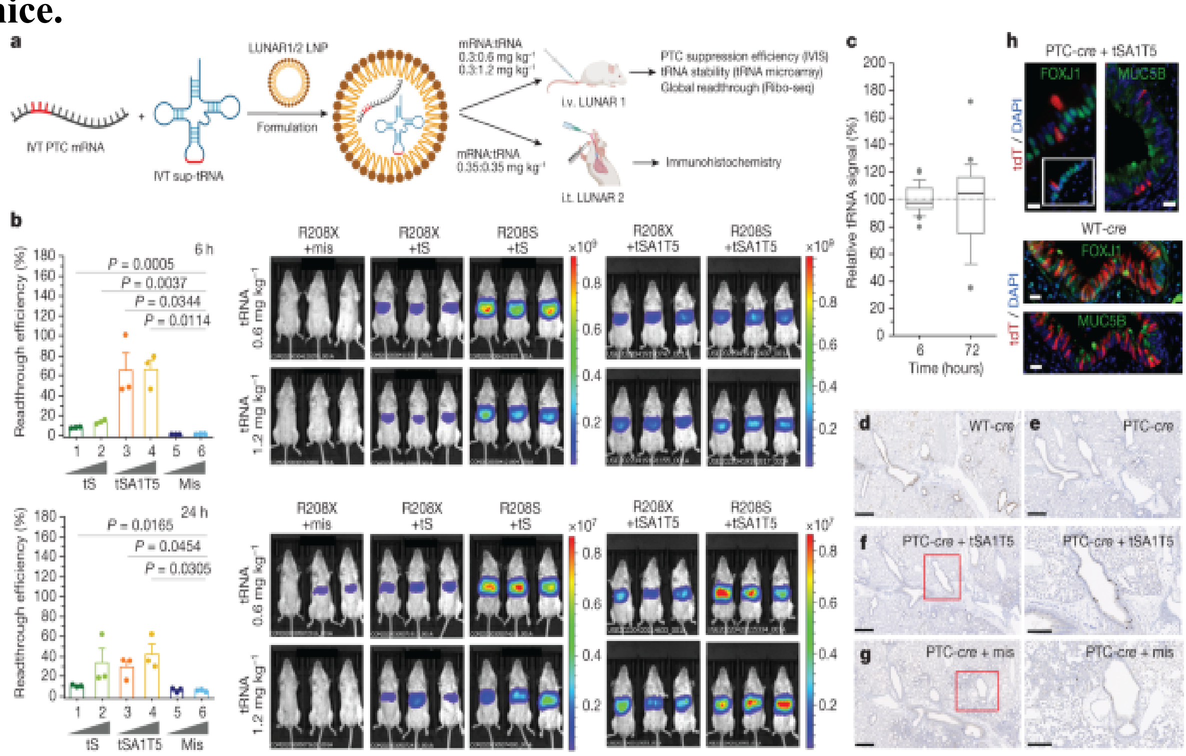
Next, we assessed whether the engineered sup-tRNA stimulates the mammalian innate immune response through activation of human Toll-like receptors (TLRs) and specifically TLR7 and TLR8, which are augmented by synthetic and viral RNA[27]. In a model system established for such analysis[28] (human TLR-transformed HEK293 cells), sup-tRNA did not activate TLR7 and only marginally activated TLR8 to the extent of the mock transfection (Fig. 1e), suggesting that it might be a non-specific effect.

In vivo efficacy of LNP–sup-tRNA

To assess the suitability of the optimized sup-tRNA as therapeutic agents, we encapsulated them in LNPs and tested the efficacy of their suppression of PTCs and their molecular safety in vivo (Fig. 2a). For intravenous administration of the tS variants in mice, we used LUNAR LNPs, which are similar to those recently established for administration of human factor IX (*F9*) mRNA in a mouse model of liver haemophilia B[21], with a total lipid-to-

RNA weight ratio of 25:1 (LUNAR_2021-1, hereafter referred to LUNAR1) (Supplementary Table 3). The LUNAR1 co-formulations of PTC akaluciferase (*aLuc*^{R208X}) mRNA (0.3 mg mRNA per kg) and either tS or tSA1T5 were delivered in two doses (0.6 and 1.2 mg sup-tRNA per kg). Six hours after intravenous administration, we detected a readthrough of up to 66% for tSA1T5 and 13% for tS (Fig. 2b), suggesting a rapid and efficient production of functional protein. Twenty-four hours after administration, the readthrough efficiency decreased to 40% (Fig. 2b), owing to the overall drop of the amount of *aLuc*^{R208X} mRNA (Extended Data Fig. 5a). By contrast, the amount of tSA1T5 remained stable for at least 72 h, as detected by tRNA-tailored microarrays (Fig. 2c and Extended Data Fig. 5b,c). The sup-tRNA stability substantially exceeds the much shorter half-life of mRNA-based therapies and vaccines[29,30].

Fig. 2: In vivo PTC suppression by LUNAR-encapsulated sup-tRNAs in mice.



a, Workflow of intravenous (i.v.) or intratracheal (i.t.) administration of LUNAR LNPs with co-encapsulated PTC reporter mRNA and sup-tRNA. For intravenous administration, *aLuc*^{R208X} (X represents UGA) reporter and two different concentrations of tS or tSA1T5 were administered in Balb/C mice. For intratracheal administration, PTC-*cre* reporter (S69X/S82X, where

X represents UGA) and tSA1T5 were administered in transgenic flox-tdTomato mice. Ribo-seq, ribosome sequencing. **b**, Quantification of the PTC readthrough (left) represented as a ratio of the signal from aLuc(R208X) mice relative to the corresponding aLuc(R208S) mice treated with sup-tRNA from IVIS in vivo images (right). Groups 1–6 were co-administered with $aLuc^{R208X}$ mRNA. Groups 1, 3, 5: 0.6 mg kg^{-1} sup-tRNA; groups 2, 4, 6: 1.2 mg kg^{-1} sup-tRNA. The mean value of the corresponding aLuc(R208S) plus sup-tRNA group is set to 100%. Data are mean ± s.e.m. ($n = 3$ per group). One-sided *t*-test. **c**, sup-tRNA stability in liver monitored with tRNA microarrays represented as a box plot (36 probes on each array, $n = 2$ independent replicates) normalized to the mean of the signal at 6 h after intravenous administration, which is set as 100% (horizontal line). The centre line indicates the median, box edges bound 10th to 90th centiles, and whiskers represent the range of the remaining data without exclusion of outliers. **d**–**g**, Bright-field immunohistochemistry showing tdTomato expression (dark brown spots) in lungs from transgenic mice treated with LUNAR2 LNPs carrying a wild-type *cre* (WT-*cre*) (**d**), PTC-*cre* (**e**) or PTC-*cre* with tSA1T5 treatment (**f**) or mismatch tRNA (**g**). In **f**,**g**, images on the right are magnified views of the outlined region in the left image. $n = 4$ mice in each group (Extended Data Fig. 7). Scale bars, 500 µm; magnified images (**f**, **g**, right), 300 µm. **h**, Immunofluorescence of tdTomato (tdT) and ciliated (FOXJ1) and secretory (MUC5B) epithelial cell markers in a mouse dosed with a LUNAR2 formulation carrying wild-type *cre* or PTC-*cre* with tSA1T5 treatment. Nuclei were counterstained with DAPI. Scale bars, 25 µm.

Source Data

To access the broader applicability of sup-tRNA for treating other nonsense mutation-linked genetic disorders whose underlying protein is specifically expressed in lung, we next administered tSA1T5 into the lungs of a flox-tdTomato transgenic mouse model by intratracheal microsprayer instillation (Fig. 2a). Here we used another lung-targeting LNP formulation with a total lipid-to-RNA weight ratio of 15:1 (LUNAR-2021-2, hereafter referred to as LUNAR2) (Supplementary Table 3). tSA1T5 was co-administered with PTC-*cre* mRNA harbouring two UGA PTCs (S69X and S82X) at two equal consecutive doses (0.35 mg kg^{-1} of each tRNA and mRNA on day 0 and day

2; Fig. 2a). We benchmarked the suitability of the PTC-*cre–loxP* model in cell culture (Extended Data Fig. 6). In the transgenic *flox*-tdTomato mice, the expression of the fluorescent tdTomato is interrupted by a *loxP*-flanked STOP cassette. External delivery of wild-type *cre* mRNA resulted in robust tdTomato fluorescence across the large and small epithelial airways (Fig. 2d and Extended Data Fig. 7). We administered PTC-*cre* mRNA that rendered *cre* mRNA unable to mediate recombination and consequently no tdTomato fluorescence was detected (Fig. 2e and Extended Data Fig. 7). LUNAR2-encapsulated co-delivery of tSA1T5 with the PTC-*cre* mRNA efficiently restored the tdTomato expression (Fig. 2f and Extended Data Fig. 7), with a pattern similar to that in cells expressing the wild-type *cre* (Fig. 2h and Extended Data Fig. 7); that is, tdTomato was expressed by two major epithelial cell populations in epithelial airways—ciliated (indicated by co-localization with forkhead box protein J1 (FOXJ1), a transcription factor regulating cilia gene expression and motile cilia formation) and secretory (indicated by co-localization with MUC5B, the major gel-forming mucin in lung, which is secreted by airway secretory cells). For the mismatch tRNA, we observed few recovery spots (Fig. 2g and Extended Data Fig. 7), probably owing to the reported low Cre-independent background tdTomato expression for this transgenic mouse.

Safety of LNP–sup-tRNA in mice

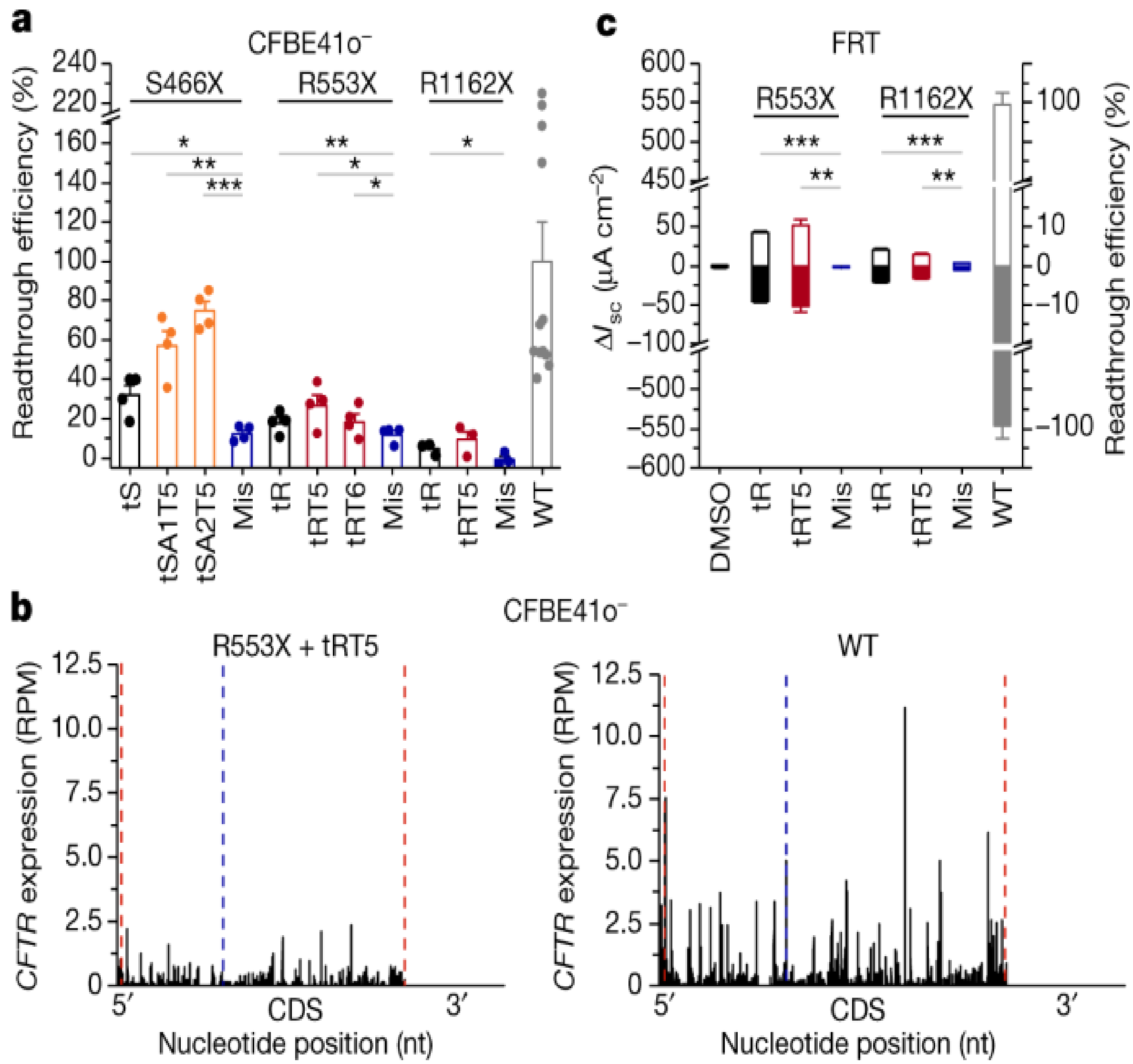
To assess potential off-target effects of sup-tRNA on native stop codons, we analysed the whole lung and liver organs of mice treated with tSA1T5 intratracheally or intravenously, respectively, using ribosome profiling. The readthrough frequency at canonical stop codons was determined using the ribosome readthrough score[26] (Methods). Out of more than 10,000 transcripts (Supplementary Table 4), we detected readthrough events at canonical UGA stop codons of a small number of transcripts (that is, 23 in the lung tissue by intratracheal administration and 15 in the liver by intravenous administration; Extended Data Fig. 8a,b); however, a similar number of transcripts underwent readthrough at UGA codons in the untreated control mice and at the other two stop codons (UAA and UAG) not targeted by the sup-tRNA (Extended Data Fig. 8a,b), indicating that tSA1T5 did not enhance the readthrough at native stop codons beyond the stochastic background readthrough level. Even at transcript-internal UGA

sites that are naturally selected for high readthrough efficiency[31], we detected no tSA1T5-triggered enhancement beyond the basal readthrough level (Extended Data Fig. 8c,d). Together, these data indicate efficacious suppression of readthrough in mice, with a good molecular safety profile, and reinforce the potential of sup-tRNA for correcting of PTC-triggered liver or respiratory diseases.

sup-tRNA efficacy on protein expression

In a pilot experiment to benchmark the efficacy of the tS and tR variants in restoring translation and expression of a full-length disease protein (that is, with intron-less cDNA), we co-transfected CFBE41o^{-} cells with in vitro-transcribed sup-tRNAs and various PTC-*CFTR* variants using Lipofectamine. The sup-tRNAs derived from each tR and tS (for example, tSA1T5, tSA2T5, tRT5 and tRT6), which showed high readthrough efficacy in in vitro screening (Fig. 1c), displayed different efficiencies in restoring full-length CFTR (that is, fully glycosylated, mature CFTR (band C)), with tS variants resulting in up to 75% and tR variants resulting in up to 27% of the expression level of cells transfected with wild-type *CFTR* (Fig. 3a and Extended Data Fig. 9a). Overall, tSA1T5 and tSA2T5 exhibited higher restoration efficacy than tRT5 and tRT6 (Fig. 3a and Extended Data Fig. 9a); this is in stark contrast to their similar readthrough activity with the reporter constructs (Fig. 1c), implying the importance of the much larger PTC sequence context on the sup-tRNA efficacy. Of note, there is a high variation in the expression levels of wild-type CFTR (Fig. 3a), which is reported to be a consequence of its complex biogenesis and the simultaneous degradation of endoplasmic reticulum-retained immature CFTR forms[32]. Using ribosome profiling, we determined that sup-tRNA uniformly restored translation, as exemplified by the effect of tRT5 on *CFTR*^{R553X}, which restored the translation level to 22% of the level of cells transfected with wild-type *CFTR* (Fig. 3b). The uniform ribosome coverage (that is, the nearly equal mean coverage of the ribosome profiling spectra) upstream and downstream of the PTC is indicative of no substantial ribosomal drop-off at the PTC (Fig. 3b).

Fig. 3: PTC suppression and restoration of mRNA translation and protein function in cell models.



a, Efficacy of tS and tR variants in restoring expression of full-length CFTR from CFTR(S466X), CFTR(R533X) or CFTR(R1162X) (where X represents the UGA codon) in CFBE41o⁻ cells, monitored by immunoblot (Extended Data Fig. 9a). Full-length CFTR (band C) expression is normalized to that in CFBE41o⁻ cells with wild-type *CFTR*. Data are mean ± s.e.m. S466X and R553X: $n = 4$; R1162X: $n = 3$; wild type: $n = 12$ independent replicates). One-sided *t*-test. **b**, Ribosome density profile of tRT5-suppressed $CFTR^{R553X}$ mRNA translation monitored by ribosome profiling (total expression, 27 reads per kilobase per million mapped reads (RPKM)) compared with wild-type *CFTR* mRNA (129 RPKM) in CFBE41o⁻ cells. Red dashed lines denote the start and stop of *CFTR* coding sequence (CDS); blue dashed line denotes the first nucleotide of the UGA

PTC. RPM, reads per million mapped reads. **c**, Short-circuit current (ΔI_{sc}) of CFTR(R553X)- and CFTR(R1162X)-expressing FRT cell monolayers transfected with tR or tRT5 compared with cells expressing wild-type CFTR. Positive values (white bars) indicate ΔI_{sc} following CFTR activation (with forskolin and VX-770) and negative (solid bars) indicate ΔI_{sc} following CFTR inhibition (with Inh-172 inhibitor). Data are mean ± s.e.m. DMSO: $n = 5$; R553X, mismatch, $n = 5$; R553X, tRT5: $n = 4$; R553X, tR: $n = 3$; R1162X, tRT5: $n = 4$; R1162X, mismatch: $n = 4$; R1162X, tR: $n = 7$; wild type: $n = 3$ independent replicates. One-sided *t*-test. For gel source data, see Supplementary Fig. 1.

Source Data

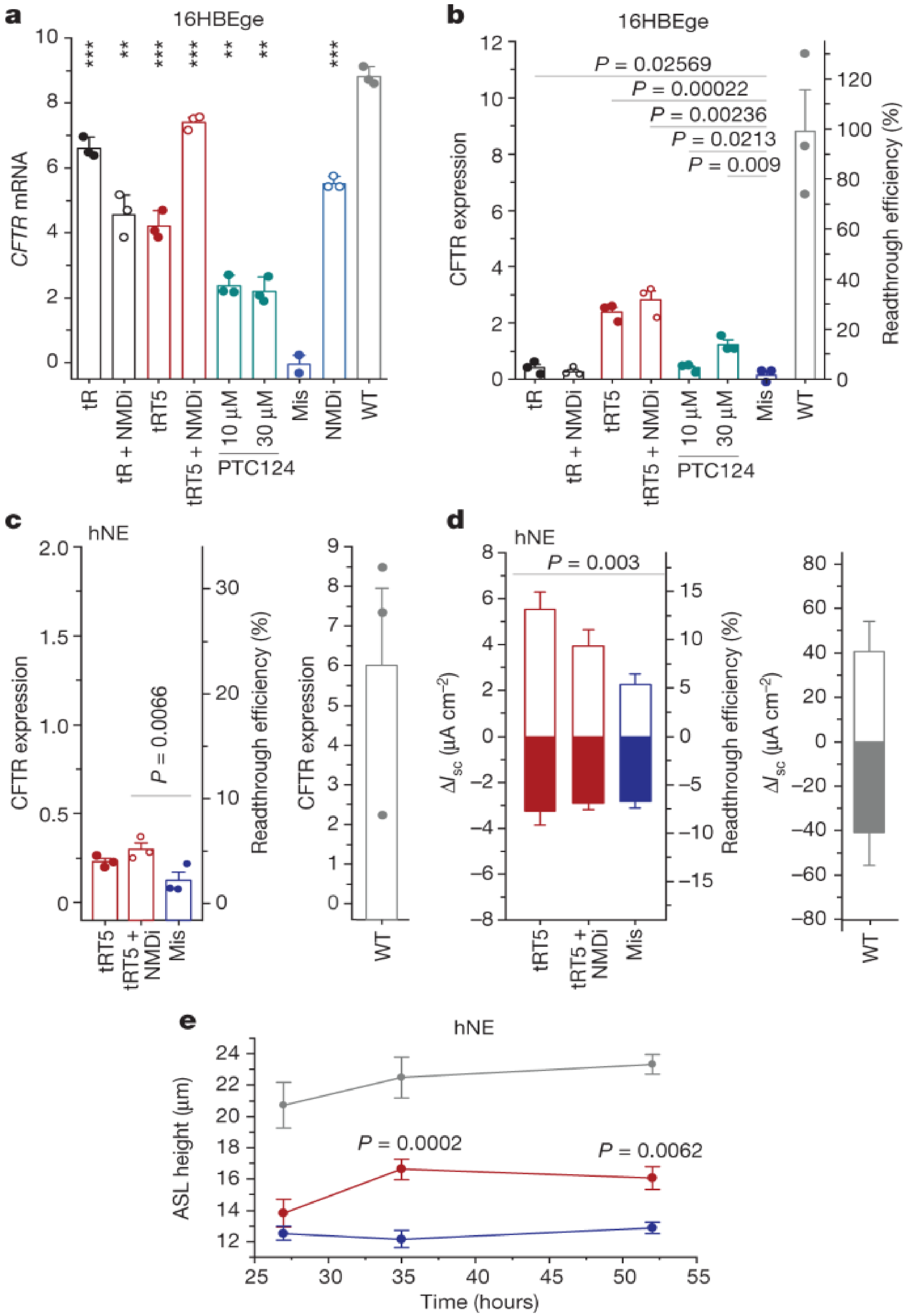
Nonsense mutations at arginine codons are the most common PTC mutations in patients with cystic fibrosis (https://cftr2.org/). Consequently, for the subsequent studies we focused on the engineered tR variants and compared their efficacies in rescuing protein expression and function of two PTC-containing *CFTR* variants, $CFTR^{R553X}$ and $CFTR^{R1162X}$. tRT5 effectively augmented the expression of $CFTR^{R553X}$ and $CFTR^{R1162X}$ by up to 27% and 10%, respectively (Fig. 3a), implying sequence context dependence of sup-tRNA efficacy. We next measured the activity of CFTR ion channels in Fischer rat thyroid (FRT) cells—a standard cellular model viewed by the US Food and Drug Administration as informative for drug label expansion of CFTR modulator compounds[25]. FRT cells were modified to stably express intron-less full-length $CFTR^{R553X}$ or $CFTR^{R1162X}$. Both tR and tRT5 restored CFTR channel activity, with similar efficacies for each PTC-CFTR variant, and augmented channel activity of CFTR(R553X) and CFTR(R1162X) by up to 9% and 3%, respectively (Fig. 3c and Extended Data Fig. 9b,c), thereby mirroring overall the protein expression levels following sup-tRNA treatment (compare Fig. 3c with Fig. 3a).

sup-tRNA antagonizes NMD

It is well documented that native transcripts containing nonsense mutations are susceptible to NMD[15,33,34], but efficient readthrough with chemical compounds or sup-tRNAs antagonize NMD on PTC-containing

mRNAs[7,26,35,36]. To assess the effect of NMD on sup-tRNA-mediated mRNA utilization, we used two systems that endogenously express full-length *CFTR*^{R1162X} (that is, with all introns and exons): (1) the gene-edited bronchial epithelial cell line 16HBEge (*CFTR*$^{R1162X/-}$; where X represents UGA), and (2) human nasal epithelial (hNE) cells obtained by non-invasive nasal brushings from patients with cystic fibrosis harbouring the homozygous nonsense mutation *CFTR*^{R1162X}, where X represents UGA. Using Lipofectamine, we transfected in vitro-transcribed tRT5 or tR into untreated 16HBEge cells or cells treated with an NMD inhibitor (5 µM NMD14) and assessed mRNA and protein expression. tR and tRT5 alone markedly stabilized the expression of full-length *CFTR* mRNA at levels similar to those achieved with the NMD inhibitor alone, with tR having a greater stabilization effect than tRT5 (Fig. 4a). A mismatched tRNA did not stabilize *CFTR* mRNA (Fig. 4a), suggesting that the effect is specific to the sup-tRNA. Combined treatment with tR or tRT5 and NMD inhibitor augmented the levels of *CFTR* mRNA, but the effect was not uniformly additive for both tRNAs (Fig. 4a), and only marginally enhanced full-length CFTR protein (band C) expression compared with tRT5 alone (Fig 4b and Extended Data Fig. 10a). For comparison, PTC124 (also known as ataluren), which is clinically approved for the treatment of Duchenne muscular dystrophy and has been shown to confer readthrough at PTCs[10,37], modestly stabilized *CFTR* mRNA and enhanced protein expression to a level similar to that of tR with and without NMD inhibitor (Fig. 4a,b). In proliferating 16HBEge cells, the transfected sup-tRNA remained stable for at least 72 h (Extended Data Fig. 5c), thus corroborating the stability observed in mice (Fig. 2c).

Fig. 4: Restoration of CFTR expression and activity by outcompeting NMD.


a
16HBEge
CFTR mRNA
***
**
***
***
**
**
***
tR
tR + NMDi
tRT5
tRT5 + NMDi
10 μM
30 μM
PTC124
Mis
NMDi
WT
b
16HBEge
CFTR expression
Readthrough efficiency (%)
P = 0.02569
P = 0.00022
P = 0.00236
P = 0.0213
P = 0.009
c
hNE
CFTR expression
Readthrough efficiency (%)
P = 0.0066
tRT5
tRT5 + NMDi
Mis
WT
d
hNE
P = 0.003
ΔIsc (μA cm−2)
Readthrough efficiency (%)
e
hNE
P = 0.0002
P = 0.0062
ASL height (μm)
Time (hours)

a,**b**, Efficacy of tR or tRT5 with or without NMD inhibitor (NMDi, 5 µM NMD14) compared with treatment with PTC124 in augmenting *CFTR* mRNA (**a**) or CFTR(R1162X) protein (**b**; band C) expression in 16HBEge cells (cells were wild type or $CFTR^{R1162X/-}$, where X represents UGA). **a**, *CFTR* mRNA level was normalized to that in untreated cells. **b**, Band C intensity was normalized to the total protein (left *y*-axis) and to wild-type CFTR (right *y*-axis). Data are mean ± s.d. of $n = 2$ independent replicates for mismatch (**a**) and mean ± s.e.m. of $n = 3$ independent replicates for all other data in **a**,**b**. One-sided *t*-test. **c**, Left, efficacy of tRT5 with and without NMD inhibitor (5 µM NMD14) for restoration of CFTR protein expression (band C is normalized to total protein (left *y*-axis)) in $CFTR^{R1162X/R1162X}$ hNE cells (X represents UGA), monitored by immunoblot. Right, CFTR expression in hNE cells from non-cystic fibrosis individuals who are wild-type for the *CFTR* gene. Data are mean ± s.e.m. ($n = 3$ independent replicates) and are shown as a percentage of the mean expression in CFTR wild-type cells. One-sided *t*-test. **d**, Left, short-circuit current measurements in $CFTR^{R1162X/R1162X}$ hNE cells (X represents UGA) with tRT5 alone or with tRT5 plus NMD inhibitor (0.5 µM SMG1) (right). Right, short-circuit currents in hNE cells from non-cystic fibrosis individuals who are wild-type for the *CFTR* gene. Positive values (white bars) indicate ΔI_{sc} following CFTR activation with forskolin and VX-770 and negative values (solid bars) indicate ΔI_{sc} following CFTR inhibition with Inh-172. Data are mean ± s.e.m. (tRT5 and mis: $n = 8$; tRT5 + NMDi and wild type: $n = 4$, independent replicates) and are shown as a percentage of the mean of CFTR expression in wild-type cells. One-sided *t*-test. **e**, ASL height measurement on $CFTR^{R1162X/R1162X}$ hNE cells (X represents UGA) co-transfected with tRT5 or mismatch tRNA, compared with CFTR expression in wild-type cells. Data are mean ± s.e.m. ($n = 5$ independent replicates). Two-way ANOVA with Sidak's multiple comparisons. For gel source data see Supplementary Fig. 1.

Source Data

sup-tRNA efficacy in hNE cells

hNE cells derived from patients with cystic fibrosis who are homozygous for the R1162X mutation were grown at an air–liquid interface, differentiated into a ciliated pseudostratified epithelial monolayer and transfected with tRT5 alone or in combination with NMD inhibitor treatment. For transfection, we used Lipofectamine, as recommended when working with isolated cells[38]. LNPs are more effective in vivo, probably because endocytosis-driven uptake of LNPs is affected by gene expression changes that occur when cells are removed from their natural environment[38]. Combined treatment with tRT5 and the NMD inhibitor NMD14 slightly increased the expression of full-length CFTR over the expression level with tRT5 alone (Fig. 4c and Extended Data Fig. 10a), corroborating our results in 16HBEge cells (Fig. 4b). In hNEs obtained by non-invasive nasal brushings from individuals without cystic fibrosis, the expression and activity of wild-type CFTR varied over a fourfold range (Fig. 4c,d); in comparisons, we used a mean value across individuals. We also noted alterations of hNE viability and wild-type CFTR expression following prolonged treatment with the NMD inhibitor (Extended Data Fig. 10b), and to exclude effects driven by the NMD inhibitor, we considered another NMD inhibitor (SMG1). tRT5 alone effectively augmented ion transport by up to 14% of the wild-type activity (Fig. 4d and Extended Data Fig. 10c), which exceeds the widely used therapeutic threshold for CFTR activity[39,40] of approximately 10%. Combined treatment with NMD inhibitor weakened the effect (Fig. 4d). Considering the intrinsically low transfection efficiency of hNE cells (approximately 20% in these studies), we acknowledge that much higher efficacy of restoration of CFTR function might be achieved in vivo. Together, our results suggest that engineered sup-tRNA alone successfully outcompetes mRNA surveillance mechanisms and that combined NMD inhibitor therapy in patients may cause adverse effects[41].

Next, we tested the ability of tRT5 to restore hNE cell function, as indicated by the thickness of the airway surface liquid (ASL)—the tightly regulated thin liquid layer that has a major role in mucus clearance and lung defence against infection and is used as a predictor of the therapeutic outcome[42]. tRT5 restored the ASL height on hNE cells after 35 h (Fig. 4e), reinforcing its potential clinical benefits.

Discussion

Here we demonstrate that native tRNAs can be modified to efficiently decode clinically important nonsense mutations. Engineering of tRNAs to decode various nonsense mutation-derived PTCs, including the most difficult to correct UAA PTC, involves altering tRNA sequences that modulate decoding accuracy[18,20] (through anticodon-stem mutations) and binding affinity to elongation factor[19] (through modulation of the TΨC-stem), framed in terms of the individual thermodynamic contribution of the amino acid to be inserted at the PTC. We showed that at PTCs that cause cystic fibrosis disease, an optimized suppressor tRNA alone restored protein expression and function, and airway volume homeostasis in a manner suggesting potential clinical benefit for individuals with cystic fibrosis. In this case, the addition of NMD inhibitors as adjuvants may not be necessary. Although NMD inhibitors may lengthen the kinetic window for PTC suppression strategies[43], the treatment and dose to block NMD must be carefully tuned to avoid widespread misregulation of gene expression[41].

This study therefore provides a framework for the development of tRNA-based PTC inhibitors for administration as lipid nanoparticles that are shaped to an individual PTC context and relevant amino acid as a means to advance the precision and individualized therapy of hereditary diseases caused by nonsense mutations.

Methods

Plasmid and RNA constructs

To test the readthrough efficiency of sup-tRNAs, we used two different constructs. First, a dual firefly luciferase–*Renilla* luciferase (*FLuc-RLuc*) reporter containing 15 codons from tripeptidyl peptidase 1 gene (codons 201–215) downstream of the AUG codon of *FLuc* (*FLuc*^{R208X}-*RLuc*). The tripeptidyl peptidase 1 (TPP1) gene associated with the autosomal recessive progressive lysosomal disorder, late infantile neuronal ceroid lipofuscinosis (CLN2). As a control, when using the tSA1T5 suppressor charged with Ser, we replaced the UGA PTC in *FLuc*^{R208X}-*RLuc* with the AGC codon encoding Ser yielding *FLuc*^{R208S}-*RLuc*. The expression of RLuc is controlled by the coxsackievirus B3 internal ribosome entry site. Second,

stretch of 15 codons (45 nt) from different disease-related genes centred at the respective PTC mutations (Supplementary Table 2) were inserted into pGL4.51 (Promega) harbouring the *luc2* gene, at the 5′ *luc2* CDS, directly after the AUG start codon, yielding PTC-*FLuc* variants.

For the in vivo experiments, the akaluciferase gene[44] (*aLuc*) was synthesized de novo (Genewiz) and cloned into pARM2379[45]. The coding sequence (CDS) of *aLuc* was extended 5′ upstream in-frame (after the *aLuc* ATG start codon) by 15 codons (45 nt) centered at disease-related PTCs, e.g. R208X of human tripeptidyl peptidase 1 gene TPP1 (codon positions 201–215) (Supplementary Table 2). R208 was mutated to R208X (UGA, UAG or UAA) to mimic the human PTC ($aLuc^{R208X}$) or to R208S ($aLuc^{R208S}$) to be used as positive control. For in-cell experiments, 15 codons flanking R208X of TPP1 and S466X (codon positions 459–473) of the human *CFTR* gene (Supplementary Table 2) were fused to *FLuc* gene yielding $FLuc^{R208X}$ and $FLuc^{S466X}$. S466 was mutated to UAA, UAG or UGA.

The *cre* gene[46] was de novo synthesized (Genewiz) extended at its 5′ end with SV40 large T-antigen nuclear localization signal (amino acid sequence PKKKRKV) to facilitate recombination efficiency[47]. Nonsense mutations at positions Ser69 and Ser82 were introduced by PCR-based mutagenesis and de novo gene synthesis with gBlock (Integrated DNA Technology).

Cell lines and primary cells

Hep3B (HB-8064) and Hepa1-6 (CRL-1830) cell lines were obtained from the ATCC. A Cre-*loxP* reporter was integrated in HEK293T (CRL-3216) cells (generated by R. Trelles). The immortalized cystic fibrosis bronchial epithelial cell line CFBE41o^{-} (generated by D. Gruenert) with no allelic CFTR expression was used for ectopic expression of CFTR variants. Transepithelial ion transport was measured in Fischer rat thyroid (FRT) cells stably expressing $CFTR^{R553X}$, $CFTR^{R1162X}$ or wild-type *CFTR*.

16HBE14o− cells, the immortalized version of human bronchial epithelial cells expressing wild-type *CFTR* (including all introns; generated by D. Gruenert) were obtained from M. Lalk[48]. 16HBE14o− cells were gene-edited at the endogenous *CFTR* locus using CRISPR–Cas9 to create isogenic

16HBEge $CFTR^{R553X/-}$ and $CFTR^{R1162X/-}$ cell lines[49]. Both 16HBEge cell lines were obtained from the Cystic Fibrosis Foundation Therapeutics Lab.

Primary hNE cells were collected by nasal brush from an individual with cystic fibrosis who was homozygous for $CFTR^{R1162X}$, and were obtained at passage 2 from the Cystic Fibrosis Foundation Therapeutics Lab.

IL-8 response was monitored in HEK293XL-TLR7 or HEK293XL-TLR8 cells (Invivogen). The cells were aliquoted at low passage number and stored in liquid nitrogen. The cells were regularly tested for mycoplasma contamination using Venor GeM PCR-based detection kit (Merck).

tRNA design

In the isoacceptors, $tRNA^{Ser}$UGA/$tRNA^{Ser}$AGA, $tRNA^{Arg}$UCU and $tRNA^{Gly}$UCC, the anticodon was exchanged to UCA produce tS, tR and tG, respectively (Supplementary Table 1). Previous studies have suggested that among the tRNA families, these isoacceptors (such as $tRNA^{Ser}$(AGA) and $tRNA^{Ser}$(UGA), $tRNA^{Arg}$(UCU) and $tRNA^{Gly}$(UCC)[4]) were among those with the highest readthrough following the exchange of their native anticodon to decode a stop codon.

For TΨC-stem tRNA variants (T_i variants), we estimated the $\Delta\Delta G°$ values for binding affinities to eEF1A considering the cumulative contribution of the three TΨC-stem base pairs 49–65, 50–64 and 51–63 (Fig. 1a, bottom right). The eukaryotic (eEF1A) and bacterial (EF-Tu) elongation factors share conserved sites of aminoacyl-tRNAs binding[50], thus, the $\Delta\Delta G°$ value for each nucleotide pair was taken from the bacterial EF-Tu–$tRNA^{Phe}$ complex[23,51].

Both anticodon-stem and anticodon-loop have coevolved with the anticodon to ensure faithful decoding. To increase the decoding, U–A or A–U pairs are preferred at nucleotide positions 31 and 39 in native sup-tRNAs[52], thus, we considered them in our A1 variants (Fig. 1a and Extended Data Fig. 1). In the A2 variants, the anticodon-stem is taken from $tRNA^{Sec}$. The identity elements—that is, nucleotides recognized by the cognate aminoacyl-tRNA synthetase to aminoacylate tRNA—were preserved. For example, in the

tRNA^{Ser} family, the discriminator base G73 and the V-region[53] act as identity elements, in the tRNA^{Arg} family they are G73, A20 and C35–U/G36, and in the tRNA^{Gly} family they are A73, C2–G71[53].

tRNA transcription

In vitro-transcribed tRNA variants were used for co-transfection in the immortalized cell culture models, patient-derived primary cells and in vivo, in mice.

tRNAs were transcribed in vitro using T7 transcription system as described[17]. In brief, two partially overlapping DNA oligonucleotides encoding the corresponding tRNA sequence with an upstream T7 promoter (5′-TAATACGACTCACTATA-3′) were used for in vitro tRNA synthesis. A 24 µM solution of both oligonucleotides was denatured for 2 min at 95 °C, and thereafter aligned for 3 min at room temperature in 20 mM Tris-HCl (pH 7.5), and 0.4 mM dNTPs were added and incubated with 4 U µl^{-1} RevertAid Reverse Transcriptase (Thermo Fisher Scientific) for 40 min at 37 °C. This dsDNA template was purified with phenol/chloroform, washed with 80% ethanol and resuspended in DEPC-treated water. Alternatively, dsDNA templates were prepared by PCR reaction on the tSA1T5-expressing plasmid with a forward primer annealing upstream of the T7 promoter and a reverse primer (5′-TGGCGTAGTCGACGGATTC-3′) with or without 2′-*O*-methyl modification of the first two nucleotides of the 5′ end[54]. For in vitro T7 transcription, 2 mM NTPs, 5 mM GMP, 1 × transcription buffer, 0.6 U µl^{-1} T7 RNA polymerase (Thermo Fisher Scientific) were added to the dsDNA template and incubated overnight at 37 °C. The tRNA variants were resolved by preparative denaturing polyacrylamide gel electrophoresis (PAGE) and eluted in 50 mM potassium acetate, 200 mM KCl pH 7.0 overnight at 4 °C, followed by ethanol precipitation and re-suspension in DEPC-treated water.

tRNAs for in vivo delivery were transcribed by T7 RNA polymerase from linearized plasmids under similar conditions to those described above and purified as previously described[21,45]. The integrity of the purified tRNAs was monitored by toluidine blue staining (0.4% (w/v)) in a 10% denaturing TBE-Urea gel (Thermo Fisher Scientific).

In vitro transcription of mRNA

To generate high-quality mRNA transcripts for the in vivo experiments or for co-transfecting in vitro-transcribed tRNA and mRNA in cell systems, we adopted methods described previously[21,45]. The mRNA transcripts were purified through a silica column (Macherey-Nagel). Dual $FLuc^{R208X}$-*RLuc* mRNA was in vitro-transcribed with unmodified nucleotides, while $aLuc^{R208X}$ and *cre* mRNAs were synthesized with UTPs fully substituted with N1-methyl-pseudouridine. The purified mRNAs were quantified by UV absorbance and their purity (% full-length) and integrity verified by Fragment Analyzer (Agilent). Transcripts were stored in RNase-free water below −60 °C until in vitro transfection or formulation with LNPs for in vivo administration.

tRNA transfection and in vitro luciferase readthrough assay

Hep3B, Hepa1-6 or CFBE41o^{-} cells were seeded in 96-well cell culture plates at 1×10^4 cells per well and grown in Dulbecco's Modified Essential Medium (DMEM, Pan Biotech or Gibco) for Hep3B and Hepa1-6 or Minimum Essential Medium (MEM, Pan Biotech) for CFBE41o^{-} cells. All medium was supplemented with 10% fetal bovine serum (FBS, Pan Biotech) and the medium for CFBE41o^{-} cells was also supplemented with 2 mM l-glutamine (Thermo Fisher Scientific). Sixteen to twenty-four hours later, Hep3B or CFBE41o^{-} cells were co-transfected in triplicate with 25 ng PTC-*FLuc* or wild-type *FLuc* plasmids and 100 ng each in vitro-transcribed tRNA variant using Lipofectamine 3000 (Thermo Fisher Scientific). After 4–6 h, medium was replaced and 24 h after transfection cells were lysed with 1× passive lysis buffer (Promega) and luciferase activity measured with luciferase assay system (Promega) and Spark microplate reader (Tecan). G418 was added to the cells at concentration 25 μg ml^{-1} and incubated for 24 h.

Twenty-four hours after seeding, Hepa1-6 cells were transfected with 12.5 ng of 1 of the 3 in vitro-transcribed reporter mRNAs ($FLuc^{R208X}$-*RLuc*, $FLuc^{R208S}$-*RLuc* or $FLuc^{R208}$-*RLuc* with Arg at position 208) together with 50 ng of in vitro-transcribed tRNA using MessengerMax (0.2 μl per well).

For in vitro dose-response experiment (Extended Data Fig. 4), 50–0.78 ng of the in vitro-transcribed tRNA was serially diluted by twofold and co-transfected into each well with the reporter mRNA (12.5 ng). To achieve similar transfection efficiencies across different dosages of PTC-pairing tRNA, an in vitro-transcribed mismatch tRNA that does not pair to the UGA PTC was used as filler so that total tRNA of 50 ng per well was always co-transfected. Cells were incubated at 37 °C overnight, rinsed with PBS and collected by adding 20 µl per well of 1 × passive lysis buffer (Promega). Luciferase activities were measured in 10 µl lysate with the Dual-Luciferase Reporter Assay kit (Promega) on Spark microplate reader (Tecan).

tRNA toxicity

CFBE41o^{-} cells were transfected with an in vitro-transcribed sup-tRNA in a concentration series from 4,000 to 7.81 ng per well and serially diluted by twofold as described above. Cell viability was determined using the CellTiter-Glo luminescent cell viability assay (Promega) according to the manufacturer's instructions.

tRNA-induced stimulation of TLR7- or TLR8-expressing cells

Human TLR-transformed HEK293 cells are an established system to analyse tRNA-induced stimulation of human TLR7 and TLR8 and the IL-8 response was monitored as described in[28]. In brief, HEK293XL cells stably transfected with hTLR7 and hTLR8 (Invivogen) were seeded in 96-well cell culture plates at 5×10^4 cells per well and cultured in DMEM (Pan Biotech) supplemented with 10% FBS (Pan Biotech) and 2 mM l-glutamine (Thermo Fisher Scientific). Twenty-four hours later, cells were transfected with in vitro-transcribed tSA2T5 (1 µg ml^{-1}) or resiquimod (R848; Invivogen, 1 µg ml^{-1}) as control activators of the TLR7–TLR8 signalling pathway[55] using Lipofectin (Thermo Fisher Scientific). After 3 h, medium was replaced and 16 h post-transfections, interleukin IL-8 in supernatants was measured using human IL-8 ELISA Kit II (BD Biosciences) according to the manufacturer's instructions.

Mass spectrometry

Hepa1-6 cells were seeded in two 10-cm dishes at 1×10^6 cells per dish. Sixteen to twenty-four hours after seeding, the cells were co-transfected with 2 µg in vitro-transcribed dual *FLuc*^{R208X}-*RLuc* mRNA or dual *FLuc*^{R208}-*RLuc* and 4 µg in vitro-transcribed tSA1T5 using MessengerMax (Thermo Fisher Scientific). Cells were incubated at 37 °C overnight and collected by trypsinization. Mock-transfected cells were used as a negative control. Cells were rinsed four times with PBS and analysed with 2D nano PRM liquid chromatography–tandem mass spectrometry (LC–MS/MS) by Jade Bio.

Formulation of the LUNAR–RNA nanoparticulate liposomes

LNPs were produced using LUNAR, a proprietary lipid nanoparticle technology platform, at Arcturus Therapeutics. The LNPs were prepared as described previously[21,56]. Appropriate volumes of lipids dissolved in ethanol at the desired ratios were mixed with an aqueous phase containing RNA using a microfluidic device, followed by downstream processing. For the encapsulation of RNA, ∼2 mg ml^{-1} RNA was dissolved in 5 mM citrate buffer (pH 3.5). The molar percentage ratio for the constituent lipids is 50% ionizable amino lipids, 7% 1,2-distearoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids), 41.5% cholesterol (Avanti Polar Lipids), and 1.5% 1,2-dimyristoyl-*sn*-glycerol, methoxypolyethylene glycol (polyethylene glycol chain molecular mass: 2,000) (NOF America). At a flow ratio of 1:3 ethanol:aqueous phases, the solutions were combined in the microfluidic device. The total lipid-to-RNA weight ratio was ∼25:1 (LUNAR1) or 15:1 (LUNAR2). The mixed material was then diluted 3 times with Tris buffer (pH 7.5) containing 50 mM NaCl and 9% sucrose after leaving the micromixer outlet, reducing the ethanol content to 6.25%. Diluted LNP formulation was concentrated and diafiltered by tangential flow filtration using hollow fibre membranes (mPES Kros membranes, Repligen) and Tris buffer (pH 7.5) containing 50 mM NaCl and 9% sucrose to remove the ethanol. Particle size and polydispersity index (PDI) were characterized using a Zen3600 (Malvern Instruments, with Zetasizer 7.1 software). Encapsulation efficiency was calculated by determining the unencapsulated RNA content by measuring fluorescence intensity (Fi) upon addition of RiboGreen (Molecular Probes) to the LNPs and comparing this value to the total fluorescence intensity (Ft) of the RNA content obtained upon lysis of the LNPs in 1% Triton X-100. The percentage encapsulation was calculated

by the ratio (Ft − Fi)/Ft × 100%). All LNPs were associated with encapsulation efficiencies of > 90%.

Mouse experiments and imaging

All mice were purpose-bred and experimentally naive at the start of the study. Mice were chosen randomly for treatment with either control or experimental conditions without blinding. Mice were housed 5 per cage in a pathogen-free environment in Innovive disposable IVC rodent caging system with a 12 h light/dark cycle, at temperature between 19–22 °C and humidity 50–60%. Ad libitum access to standard diet (2018, Global 18% protein rodent diet from Envigo+++) and pre-filled acidified water from Innovive (pH 2.5–3.0) were used throughout the study period. The bedding material was hardwood chips (Sani-Chips, 7115, Envigo++++) and cages were changed biweekly. All in vivo procedures involving animals were performed at Arcturus Therapeutics in accordance with the animal use protocols and policies approved by the Institutional Animal Care and Use Committee (IACUC), protocol (EB17-004-003 from 1 February 2017 and latest amendment from 17 June 2021). The vivarium is managed by an AAALAC approved vendor Explora BioLabs (A Charles River Company).

Intravenous administration and in vivo imaging: 8–10 week old, female Balb/C mice were purchased from Charles River Laboratories. LUNAR1 formulations were administered intravenously at 0.9 or 1.5 mg kg^{-1} (that is, 0.3 mg kg^{-1} luciferase mRNA and 0.6 or 1.2 mg kg^{-1} sup-tRNA) on day 0. Six and twenty-four hours after dosing, Akalumine-HCL (TokeOni, Sigma Aldrich, 808350-100MG, lot no. MKCL1624, 15 mg ml^{-1}) was injected intraperitoneally at 100 µl per mouse followed by in vivo imaging. Ten minutes after administration of the luminescent Luc substrate, live animal bioluminescence imaging was performed on mice anaesthetized with 2% isoflurane using IVIS Lumina III (Perkin Elmer). Images were quantified by the region of interest for total FLuc signal using Living Image Software (Perkin Elmer). In total, 33 mice were subjected to intravenous administrations. Naive animals within the same age group were stratified to different treatment groups; animals were assigned to groups randomly. No statistical tests were used to predetermine the sample size. Six mice were dosed with LUNAR1 co-formulated with in vitro-transcribed Luc^{R208X}

mRNA and in vitro-transcribed tS at two different concentrations (each cohort comprising three mice). Six mice were dosed with LUNAR1 co-formulated with mRNA without a PTC (*Luc*^{R208S}) and tS at two different concentrations (each cohort comprising three mice). Another cohort of six mice was dosed with LUNAR1 co-formulated with in vitro-transcribed *Luc*^{R208X} mRNA and in vitro-transcribed tSA1T5 at two different concentrations (each cohort of three mice). Six mice were dosed with LUNAR1 co-formulated with mRNA without a PTC (*Luc*^{R208S}) and tSA1T5 at two different concentrations (each cohort of three mice). To another cohort of six mice LUNAR1 co-formulated with in vitro-transcribed *Luc*^{R208X} mRNA and in vitro-transcribed mismatch tRNA not pairing to the UGA PTC at two different concentrations (each cohort comprising three mice) was administered. Three mice were treated with PBS and served as negative control.

Intratracheal administration: six- to ten-week-old, female B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J mice were purchased from Jackson Laboratories. LUNAR2 formulations were administered through intratracheal instillation in mice anaesthetized with 2% isoflurane, using a 27G × 1 inch blunt needle (B27-100, SAI) at 0.35 mg kg^{-1} for wild-type *cre* mRNA only or 0.7 mg kg^{-1} of in vitro-transcribed PTC-*cre* mRNA and in vitro-transcribed tSA1T5 (0.35 mg kg^{-1} each); 50 µl per mouse was given on day 0 and day 2 (two doses total). Forty-eight hours after the last dose, lungs from anaesthetized mice were collected and insufflated with 10% neutral buffered formalin (NBF). Lungs were then fixed overnight in 10% NBF prior to processing for immunohistochemistry. In total, 24 mice were subjected to intratracheal administrations and animals were assigned to groups randomly. No statistical tests were used to predetermine sample size. Four mice were dosed with LUNAR2 co-formulated with PTC-*cre* mRNA and tSA1T5 with anticodon pairing to UGA, four with LUNAR2 co-formulated with PTC-*cre* mRNA and mismatch tSA1T5 with anticodon pairing to UAG, four with LUNAR2 co-formulated with PTC-*cre* mRNA only, four with LUNAR co-formulated with wild-type *cre* mRNA only, and four received PBS as a negative control.

Sections were cut at 5 µm, dewaxed and rehydrated into distilled water before antigen retrieval procedure. Sections were incubated with TrueBlack

Lipofuscin Autofluorescence Quencher (Biotium) for 1 min, washed in PBS and blocked to suppress non-specific antibody biding using TrueBlack IF Background Suppressor System (Biotium). Sections were incubated with primary rabbit polyclonal RFP antibody (dilution 1:800; 600-401-379, Rockland Antibody), mouse monoclonal FOXJ1 antibody (dilution 1:1,000; 14-9965-82, Thermo Fisher Scientific) or rabbit polyclonal MUC5B antibody (dilution 1:1,000; PA5-82342, Thermo Fisher Scientific) overnight at 4 °C. A M.O.M. kit (Vector, BMK-2002) was used to remove any mouse-on-mouse Ig interference. After washing in PBS, donkey anti mouse AF488 (A-21202, Thermo Fisher Scientific) or donkey anti rabbit AF555 (A-21428, Thermo Fisher Scientific) secondary antibodies were incubated with slides for 1 h. Slides were subsequently counterstained in DAPI and mounted with VECTASHIELD Vibrance Antifade Mounting Medium (H-1700, Vector Laboratories).

In vitro Cre-*loxP* system

Stable expression system of Cre recombinase-dependent eGFP was established by transfecting HEK293T cells with pLV-CMV-LoxP-DsRed-LoxP-eGFP[57] (a gift from J. Rheenen), with puromycin selection following a protocol from the Genetic Perturbation Platform at Broad Institute. The selected cells were confirmed with lack of eGFP by EVOS Cell Imaging Systems (Thermo Fisher Scientific). Cells were seeded in 96-well plates at 1.6×10^4 cells per well and after 24 h transfected with in vitro-transcribed *cre* mRNA (12.5 ng) and in vitro-transcribed tS variants (25 ng) pairing to different PTCs (tS::UGA, tS::UAG; tS::UAA) using MessengerMax (0.2 µl per well). Every day within four days after transfection the fluorescence was recorded on Spark microplate reader (Tecan) and reported as relative fluorescence units. The expression of DsRed and eGFP were visualized by EVOS Cell Imaging Systems (Thermo Fisher Scientific).

Culturing of patient-derived nasal epithelial cells at air–liquid interface

hNE (R1162X/R1162X) cells were seeded in T75 flasks coated with collagen IV (Sigma) or Purecol (Advanced Biomatrix) in 12 ml pre-warmed complete PneumaCult ALI Ex^+ medium (StemCell kit) at 37 °C in 5% CO_2.

To 500 ml medium the following supplements were added: 0.5 ml hydrocortisone (StemCell), 10 ml 50X Ex^+ supplement (StemCell kit), and for some preparations 2 ml of amphotericin B (12.5 μg ml^{-1}; Sigma), 500 μl ceftazidime (100 mg ml^{-1}; Sigma), 500 μl vancomycin (100 mg ml^{-1}; Sigma), and 500 μl tobramycin (100 mg ml^{-1}; Sigma). Cells were expanded for 3–5 days, until they reached 70–80% confluency.

Cells were then detached by 0.05% trypsin-EDTA (Pan Biotech) or enzymatic (StemCell) treatment and seeded onto 12- or 24-ALI Transwells (0.4-μm pore polyethylene terephtalate membrane inserts, Corning) coated with collagen IV at a confluency of 1.5×10^5 to 2×10^5 per well, and Complete Ex^+ medium (without antibiotics) was added as following, 0.5–0.6 ml (basolaterally) and ~0.5 ml (apically). Cells were grown for 3–4 days at 37 °C with 5% CO_2. Medium were changed every day on the apical and basolateral side. On day 4, the apical medium was removed, and basolateral bathing solution exchanged for ALI complete medium (StemCell), followed by additional exchanges three times per week for at least 21 days until reaching a fully differentiated state.

Transfection with tRNA and treatment with NMD inhibitors

CFBE41o$^-$ cells, 16HBEge *CFTR*$^{R1162X/-}$ or 16HBE41o− *CFTR*^{WT} cells were seeded on 12-well cell culture plates at 1×10^5 cells per well and cultured in Minimum Essential Medium (MEM, Pan Biotech) supplemented with 10% FBS (Pan Biotech) and 2 mM l-glutamine (Thermo Fisher Scientific). The medium of 16HBEge cells was additionally supplemented with 1% penicillin/streptomycin (Gibco) and the culture plates were precoated with 1% human fibronectin (Thermo Fisher Scientific), 1% bovine collagen type I (Advanced BioMatrix), 1% bovine serum albumin (Gibco). At 24 h after seeding, CFBE41o$^-$ cells were co-transfected with 400–500 ng PTC-*CFTR* (S466X, R553X or R1162X variants) or wild-type *CFTR* plasmids and 800–1,000 ng in vitro-transcribed tRNA variant using Lipofectamine 3000 (Thermo Fisher Scientific). In the experiments with the NMD inhibitor (immunoblot and quantitative PCR with reverse transcription (RT–qPCR) analysis), 16HBEge cells were pre-treated with 5 μM NMD14 (MedChemExpress) for 24 h and then transfected with 800 ng of in vitro-

transcribed tRNA variants using Lipofectamine 3000 (Thermo Fisher Scientific). After 4–6 h of tRNA transfection, medium was replaced and cells grown for another 24 h.

Well-differentiated hNE (R1162X/R1162X) cells (see above) were transfected from the apical surface of monolayers with 800–1,000 ng in vitro-transcribed tRNA (tRT5 or mismatch tRNA) using Lipofectamine 3000 (Thermo Fisher Scientific). After 6 h, fresh medium was exchanged and cells were incubated for the next 24 h. Transfection efficiency with Lipofectamine was approximately 18–20% as estimated using co-transfection with fluorescent proteins. In the experiments with NMD inhibitor (NMD14 or SMG1[58,59]), drug was added on the basolateral side. Six hours following addition of the NMD inhibitor, cells were transfected from the apical surface of monolayers with in vitro-transcribed tRNA, while continuing the treatment with the NMD inhibitor. After an additional 6 h (total treatment with NMD14 or SMG1 was 12 h), the medium was replaced and cells grown for additional 24 h. We benchmarked the SMG1 concentration of 0.5 μM to allow mRNA stabilization as comparable to NMD14 at 5 μM. Two inhibitors were used to exclude an inhibitor-specific effect. It should be noted that 16HBEge cells were robust to NMD treatments, whereas donor-derived hNEs exhibited alterations in viability after extended treatment with both NMD14 or SMG1, thus, NMD treatment should not exceed 12–15 h.

CFTR immunoblot expression analysis

Cells (CFBE41o^{-} cells, 16HBEge $CFTR^{R1162X/-}$, 16HBE41o− $CFTR^{WT}$ or hNE $CFTR^{\mathrm{R1162X/R1162X}}$) transfected with tRNA or mock-treated, were then lysed with 80 μl MNT buffer (10×; 300 mM Tris-HCl pH 7.5, 200 mM MES and 1 M NaCl) and lysates were subjected to immunoblotting with monoclonal CFTR-NBD2 antibody (1:100 dilution, 596, J. R. Riordan and T. Jensen) available through the Cystic Fibrosis Foundation Therapeutics Antibody Distribution Program. Analysis utilized the capillary electrophoresis system (Jess, ProteinSimple) as described previously[60]. The peak area corresponding to fully glycosylated CFTR (band C) was normalized to the total protein with the Jess quantification module and to the molecular weight marker (180 kDa peak) to minimize fluctuations between

multiple Jess runs. Note that mock-transfected cells—those treated with Lipofectamine—were used as a control due to the slight adverse effect of Lipofectamine. Ataluren (PTC124) was added at a final concentration of 10 μM and 30 μM and incubated for 24 h.

RT–qPCR

The steady-state mRNA expression of *CFTR* variants transfected with in vitro-transcribed tRNA variants (1,000 ng) or NMD inhibitor (5 μM) was measured using RT–qPCR. Cells were grown and treated the same way as described above for the immunoblot analysis, but to capture the mRNA expression, cells were lysed 6 h after the tRNA transfection (total treatment with NMD14 and tRNA was 12 h and 6 h, respectively). Total RNA was isolated using TRIzol (Invitrogen). 2 μg total RNA was reverse transcribed using random hexamers (Thermo Fisher Scientific) and RevertAid H Minus reverse transcriptase (Thermo Fisher Scientific) in 20 μl total volume. Quantitative PCR was performed using SensiMix SYBR Hi-ROX Kit (Thermo Fisher Scientific) on T Professional thermocycler (Biometra). The region spanning exon 23-exon 24 (nucleotides 3874–4001 in the *CFTR* mRNA) of the *CFTR* transcript was amplified with the following primer pair: forward 5′-GATCGATGGTGTGTCTTGGGA-3′ and reverse 5′-TCCACTGTTCATAGGGATCCAA-3′. *GUSB* transcript was used as a house-keeping expression control whose expression level ranges at the level of $CFTR^{WT}$ expression and was amplified using the following primer pair: forward 5′-GACACGCTAGAGCATGAGGG-3′ and reverse 5′-GGGTGAGTGTGTTGTTGATGG-3′. The analysis was performed using $\Delta\Delta C_T$ approach. Technical duplicates of each biological replicate reaction were carried out for each sample.

tRNA stability and tRNA-tailored microarrays

16HBEge $CFTR^{R553X/-}$ cells were seeded at 6×10^5 cells per well into precoated 6-well cell culture plates and grown as described above. 24 h after seeding, cells were transfected with 1.5 μg tSA2T5 or water (as control) in triplicates using Lipofectamine 3000 according to the manufacturer’s protocol. Cells were lysed at 5, 24, 36, 48 and 72 h post transfection, by adding 1 ml TRIzol per well (Invitrogen). Total RNA from cells was isolated

using the TRIzol method according to manufacturer's instructions and RNA integrity was assessed by 10% denaturing polyacrylamide gel electrophoresis.

Four female Balb/C mice were intravenously dosed with LUNAR1 formulated with 0.6 mg kg^{-1} in vitro-transcribed tSA1T5. At 6 h and 72 h post treatment, mice were anaesthetized with 3% isoflurane in a VetEquip inhalation anaesthesia system chamber. Livers from the anaesthetized mice were collected and flash frozen. Mice treated with PBS served as a negative control. The whole organs were pulverized in liquid nitrogen and lysed by grinding in 500 µl TRIzol (Invitrogen). Total RNA was isolated using the TRIzol method according to manufacturer's instructions and RNA integrity was assessed on 10% denaturing polyacrylamide gel electrophoresis.

tRNAs were analysed using tRNA-tailored microarrays as previously described[61,62], with some adjustments to measure the sup-tRNA. On the microarrays, tDNA probes covering the full-length tRNA sequence of the 41 cytoplasmic tRNA species complementary to 49 nuclear-encoding tRNA families are spotted, along with the tDNA complementary to tSA1T5 (5′-TGGCGTAGTCGACGGGATTCGAACCCGTGCGGGGAAACCCCAATGGTTTTGAAGACCATCGCCTTAACCACTCGGCCACGACTAC-3′) or tDNA complementary to tSA2T5 (5′-TGGCGTAGTCGACGGGATTCGAACCCGTGCGGGGAAACCCCAACAGGTTTGAAGCCTGCCGCCTTAACCACTCGGCCACGACTAC-3′). Each microarray consisted of 12 identical blocks, each containing 2 probes for each natural tRNA and 3 probes for tSA2T5 or tSA1T5 (36 signals in total for each sup-tRNA). To fully deacylate tRNAs, 5 µg of total RNA was incubated for 45 min at 37 °C in 100 mM Tris-HCl buffer (pH 9.0), followed by purification by precipitation with ethanol and 0.1 volume of 3 M sodium acetate (pH 5.5), supplemented with glycogen (20 mg ml^{-1}, Thermo Fisher Scientific). Cy3-labeled RNA:DNA hairpin oligonucleotide was ligated to deacylated 3′-NCCA ends of the tRNAs using T4 DNA ligase (NEB) for 1 h at room temperature. Total RNA from non-transfected cells was used as comparison and labelled with Atto647-labelled RNA:DNA hairpin oligonucleotide. For subsequent normalization of the arrays, each sample was spiked in with three in vitro-transcribed tRNAs (2 µM of each), which do not cross hybridize with any of the human tRNAs or the sup-tRNA.

Detailed experimental protocol for tRNA microarrays is available at protocols.io (https://doi.org/10.17504/protocols.io.hetb3en). Scanned microarray slides were analysed using inhouse Python scripts. The median of the ratio of Cy3 to Atto647 signals was normalized to spike-ins whose ratio set to one. Thereafter, each single Cy3 signal from the sup-tRNA was normalized to the Cy3 signal of the spike-ins and represented as a ratio to the mean of the signal at 5 h (for 16HBEge cells) or 6 h (for mouse) which was set as 100%. The arrays were performed in two biological replicates for the samples withdrawn at 5 h, 24 h and 36 h post transfection, and in a single replicate for 48 h and 72 h samples. Due to a high reproducibility of the arrays (confidence intervals higher than 98%), following the normalization to the spike-ins the individual signals from the biological duplicates were merged.

Short-circuit current I_{sc} measurements

Transepithelial ion transport was measured in FRT cells, which represent a standard model for polarizing epithelia expressing apical CFTR and are viewed by the US Food and Drug Administration as informative for drug label expansion for CFTR modulators[63,64]. FRT cells stably expressing *CFTR*^{R553X}, *CFTR*^{R1162X} or wild-type *CFTR* were seeded at 100,000–150,000 cells onto permeable supports (0.33 cm^2 per Transwell insert). Four days after seeding, cells form tight junctions and were transfected with 400 ng in vitro-transcribed tRNA (tR, tRT5 or mismatch tRNA) in 20 µl OptiMEM per insert using Lipofectamine 3000 (Thermo Fisher Scientific). Cells were maintained under air–liquid interface conditions at 37 °C in 5% CO_2 for 24 h. As described above, hNE (R1162X/R1162X) cells were also cultured and treated with NMD inhibitor as indicated, followed by I_{sc} measurement.

Short-circuit current was monitored under voltage clamp conditions with an MC8 voltage clamp and P2300 Ussing chamber equipment (Physiologic Instruments). Cells grown on culture inserts (Corning) were bathed on both sides with identical Ringer's solutions containing (in mM): 115 NaCl, 25 $NaHCO_3$, 2.4 KH_2PO_4, 1.24 K_2HPO_4, 1.2 $CaCl_2$, 1.2 $MgCl_2$, and 10 d-glucose (pH 7.4). Solutions were aerated with 95% O_2:5% CO_2, and 1-s-

long, 3-mV pulses imposed every 10 s to calculate resistance by Ohm's law. As indicated for the particular study, mucosal solutions were changed to a low chloride buffer (1.2 mM NaCl and 115 mM sodium gluconate, with other components as above). Amiloride (100µM) was added (bilaterally) to block residual sodium current, followed by the CFTR agonist forskolin (10 µM) and CFTR potentiator VX-770 (5 µM). At the end of each experiment, Inh-172 (10 µM, apically) was employed to block CFTR-dependent I_{sc}. For analysis of Ussing chamber data, the ACQUIRE & ANALYZE 2.3 package (Physiologic Instruments) was run on Windows environment software to measure current, voltage, conductance and resistance from 1 to 8 tissues simultaneously. For hNE measurements, a standard setting provided with the equipment software was used with 60 s data acquisition to monitor current changes from baseline.

Functional assessment of airway surface by micro-optical coherence tomography

For assessment of the functional microanatomic parameters (such as ASL height) of hNE (R1162X/R1162X) transfected with tRT5 or mismatch tRNA, we used micro-optical coherence tomography (µOCT), a high-speed, high-resolution microscopic reflectance imaging approach, as described earlier[65,66]. In brief, this is a non-invasive method, without using exogenous dyes and particles, to image airway epithelia and the associated quantitative analysis, and the µOCT instrument provides cross-sectional images of the cell monolayers at a resolution of approximately 1 µm. Images were acquired 1 mm from the filter periphery with a scanning beam parallel to the tangent of the circumference of the filter membrane disc. Data were acquired at 20,480 Hz line rate, resulting in 40 frames per second at 512 lines per frame. Quantification of the ASL height was performed directly by geometric measurement of the corresponding layers by an investigator blinded to treatment using Image J software. Statistical analysis was performed by two-way ANOVA using Sidak's multiple comparisons.

Ribosome profiling and data analysis

For ribosome profiling, 2 female wild-type mice were intravenously dosed with LUNAR1 formulated with 0.6 mg kg^{-1} in vitro-transcribed tSA1T5,

and 2 female mice were administered intratracheally twice 48 h apart (on day 0 and day 2) with LUNAR2 formulated with 0.35 mg kg^{-1} in vitro-transcribed tSA1T5. Six hours after treatment the mice were anaesthetized with 3% isoflurane in a VetEquip inhalation anaesthesia system chamber. With mice still under anaesthesia inhaled through a nose cone, thoracotomy was performed for dissecting liver and lung tissues, and the tissues were immediately flash frozen. Mice treated with PBS for the same duration served as negative control. The whole organs were pulverized in liquid nitrogen and lysed by grinding in 360 µl 10 mM Tris-HCl (pH 7.4) supplemented with 5 mM $MgCl_2$, 100 mM KCl, 1% NP-40, 2 mM DTT, 100 µg ml^{-1} cycloheximide topped with 40 µl 10% sodium deoxycholate. Lysate from each animal organ were used to produce an independent library.

Twenty-million CFBE41o^{-} cells were co-transfected with 400 ng $CFTR^{R553X}$ plasmid and 400 ng in vitro-transcribed tRT5 or with 400 ng wild-type CFTR plasmid alone using Lipofectamine 3000 (Thermo Fisher Scientific). Twenty-four hours after transfection, cells were collected and lysed with lysis buffer (10 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$, 100 mM KCl, 1% NP-40, 2 mM DTT).

After lysis, the lysates from mice organs or CFBE41o^{-} cells were supplemented with cycloheximide (100 µg ml^{-1}) to additionally stabilize ribosome–mRNA complexes during RNase I digestion (1.5 µl of 5 U per OD_{260} for 30 min). Sequencing libraries from the RNase I digestion-derived ribosome-protected fragments (RPFs) were prepared using a protocol for micro RNA with direct ligation of the adapters[67].

Sequenced reads were quality selected using the fastx-toolkit (0.0.13.2) with a threshold of 20. Adapter sequences were removed by cutadapt (1.8.3) with a minimal overlap of 1 nt. The libraries were depleted of reads mapping to rRNA reference sequences (bowtie 1.2.2; -y –un) and the reads were mapped to the human (GRCh38) and mouse (GRCm38) reference genomes, respectively. Mapping was performed using STAR[68] (2.5.4b) allowing maximum of one mismatch and filtering out reads mapping to multiple positions (–outFilterMismatchNmax 1–outFilterMultimapNmax 1). In the reference annotation files, the longest annotated CDS for each transcript was selected. For two transcripts of the same CDS length, we selected the longest

transcript including 5′ and 3′ untranslated regions. Uniquely mapped reads were normalized to RPM or RPKM.

To evaluate the stop codon readthrough in the mouse libraries, we used the procedure described in ref. [29]. We plotted the middle nucleotide of RPFs or by odd-length reads the nucleotide upstream to the middle in the regions flanking the stop codon, 100 nt upstream and 100 nt downstream of the stop codon. Reads with a length of 25–32 nt were considered. For the CFBE41o^{-} cell culture libraries, we first calibrated the RPFs to the A-site codon in each RPF following the described procedure along the scripts therein[69]. In these analyses, transcripts with expression higher than 0.1 RPKM were considered.

To select transcripts that have undergone readthrough, we used the ribosome readthrough score (RRTS) described in[26]. RRTS is a ratio of the mean read density over the CDS (reads per kilobases (RPK), normalized to the CDS length) and the mean read density between the natural termination codon and next in-frame stop codon (RPK, normalized to the length of the considered 3′ untranslated region between two stop codons) separated by at least 4 nt from the natural stop codon of the CDSs. The CFTR transcript (ENSEMBL: ENSG00000001626; ENST00000003084.11) coverage is represented as RPM.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data from the Ribo-seq and tRNA microarrays were deposited in the Gene Expression Omnibus (GEO) under the accession numbers GSE191048, GSE192623 and GSE205660. Source data are provided with this paper.

References

1. Mort, M., Ivanov, D., Cooper, D. N. & Chuzhanova, N. A. A meta-analysis of nonsense mutations causing human genetic disease. *Hum. Mutat.* **29**, 1037–1047 (2008).

2. Bordeira-Carrico, R. et al. Rescue of wild-type E-cadherin expression from nonsense-mutated cancer cells by a suppressor-tRNA. *Eur. J. Hum. Genet.* **22**, 1085–1092 (2014).

3. Chang, J. C., Temple, G. F., Trecartin, R. F. & Kan, Y. W. Suppression of the nonsense mutation in homozygous β^0 thalassaemia. *Nature* **281**, 602–603 (1979).

4. Lueck, J. D. et al. Engineered transfer RNAs for suppression of premature termination codons. *Nat. Commun.* **10**, 822 (2019).

5. Porter, J. J., Heil, C. S. & Lueck, J. D. Therapeutic promise of engineered nonsense suppressor tRNAs. *Wiley Interdiscip. Rev. RNA* **12**, e1641 (2021).

6. Temple, G. F., Dozy, A. M., Roy, K. L. & Kan, Y. W. Construction of a functional human suppressor tRNA gene: an approach to gene therapy for *β*-thalassaemia. *Nature* **296**, 537–540 (1982).

7. Wang, J. et al. AAV-delivered suppressor tRNA overcomes a nonsense mutation in mice. *Nature* **604**, 343–348 (2022).

8. Borgatti, M., Altamura, E., Salvatori, F., D'Aversa, E. & Altamura, N. Screening readthrough compounds to suppress nonsense mutations: possible application to β-thalassemia. *J. Clin. Med.* **9**, 289 (2020).

9. Sharma, J. et al. A small molecule that induces translational readthrough of CFTR nonsense mutations by eRF1 depletion. *Nat. Commun.* **12**, 4358 (2021).

10. Roy, B. et al. Ataluren stimulates ribosomal selection of near-cognate tRNAs to promote nonsense suppression. *Proc. Natl Acad. Sci. USA* **113**, 12508–12513 (2016).

11. Westhof, E., Thornlow, B., Chan, P. P. & Lowe, T. M. Eukaryotic tRNA sequences present conserved and amino acid-specific structural signatures. *Nucleic Acids Res.* **50**, 4100–4112 (2022).

12. Lawson, M. R. et al. Mechanisms that ensure speed and fidelity in eukaryotic translation termination. *Science* **373**, 876–882 (2021).

13. Brown, A., Shao, S., Murray, J., Hegde, R. S. & Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493–496 (2015).

14. Joazeiro, C. A. P. Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.* **20**, 368–383 (2019).

15. Karousis, E. D. & Muhlemann, O. Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harb. Perspect. Biol.* **11**, a032862 (2019).

16. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* **19**, 526–541 (2018).

17. Albers, S. et al. Repurposing tRNAs for nonsense suppression. *Nat. Commun.* **12**, 3850 (2021).

18. Nguyen, H. A., Sunita, S. & Dunham, C. M. Disruption of evolutionarily correlated tRNA elements impairs accurate decoding. *Proc. Natl Acad. Sci. USA* **117**, 16333–16338 (2020).

19. Schrader, J. M., Chapman, S. J. & Uhlenbeck, O. C. Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc. Natl Acad. Sci. USA* **108**, 5215–5220 (2011).

20. Yarus, M., Cline, S., Raftery, L., Wier, P. & Bradley, D. The translational efficiency of tRNA is a property of the anticodon arm. *J. Biol. Chem.* **261**, 10496–10505 (1986).

21. Ramaswamy, S. et al. Systemic delivery of factor IX messenger RNA for protein replacement therapy. *Proc. Natl Acad. Sci. USA* **114**, E1941–E1950 (2017).

22. Grosjean, H. & Westhof, E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* **44**, 8020–8040 (2016).

23. Uhlenbeck, O. C. & Schrader, J. M. Evolutionary tuning impacts the design of bacterial tRNAs for the incorporation of unnatural amino acids by ribosomes. *Curr. Opin. Chem. Biol.* **46**, 138–145 (2018).

24. McCague, A. F. et al. Correlating cystic fibrosis transmembrane conductance regulator function with clinical features to inform precision treatment of cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **199**, 1116–1126 (2019).

25. Costa, E., Girotti, S., Pauro, F., Leufkens, H. G. M. & Cipolli, M. The impact of FDA and EMA regulatory decision-making process on the access to CFTR modulators for the treatment of cystic fibrosis. *Orphanet J. Rare Dis.* **17**, 188 (2022).

26. Wangen, J. R. & Green, R. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. *eLife* **9**, e52611 (2020).

27. Kariko, K. & Weissman, D. Naturally occurring nucleoside modifications suppress the immunostimulatory activity of RNA: implication for therapeutic RNA development. *Curr. Opin. Drug Discov. Dev.* **10**, 523–532 (2007).

28. Kariko, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23**, 165–175 (2005).

29. Leppek, K. et al. Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat. Commun.* **13**, 1536 (2022).

30. Sahin, U., Kariko, K. & Tureci, O. mRNA-based therapeutics–developing a new class of drugs. *Nat. Rev. Drug Discov.* **13**, 759–780 (2014).

31. Loughran, G. et al. Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. *J. Biol. Chem.* **293**, 4434–4444 (2018).

32. Ward, C. L., Omura, S. & Kopito, R. R. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83**, 121–127 (1995).

33. Karousis, E. D., Gurzeler, L. A., Annibaldis, G., Dreos, R. & Muhlemann, O. Human NMD ensues independently of stable ribosome stalling. *Nat. Commun.* **11**, 4134 (2020).

34. Kurosaki, T., Popp, M. W. & Maquat, L. E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.* **20**, 406–420 (2019).

35. Ko, W., Porter, J. J., Sipple, M. T., Edwards, K. M. & Lueck, J. D. Efficient suppression of endogenous CFTR nonsense mutations using anticodon-engineered transfer RNAs. *Mol. Ther. Nucleic Acids* **28**, 685–701 (2022).

36. Keeling, K. M. et al. Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in S. cerevisiae. *RNA* **10**, 691–703 (2004).

37. McDonald, C. M. et al. Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **390**, 1489–1498 (2017).

38. Zhang, J., Shrivastava, S., Cleveland, R. O. & Rabbitts, T. H. Lipid–mRNA nanoparticle designed to enhance intracellular delivery mediated by shock waves. *ACS Appl. Mater. Interfaces* **11**, 10481–10491 (2019).

39. Masvidal, L. et al. Assessing the residual *CFTR* gene expression in human nasal epithelium cells bearing *CFTR* splicing mutations causing cystic fibrosis. *Eur. J. Hum. Genet.* **22**, 784–791 (2014).

40. Rowe, S. M., Accurso, F. & Clancy, J. P. Detection of cystic fibrosis transmembrane conductance regulator activity in early-phase clinical trials. *Proc. Am. Thorac. Soc.* **4**, 387–398 (2007).

41. Supek, F., Lehner, B. & Lindeboom, R. G. H. To NMD or not to NMD: nonsense-mediated mRNA decay in cancer and other genetic diseases. *Trends Genet.* **37**, 657–668 (2021).

42. Haq, I. J., Gray, M. A., Garnett, J. P., Ward, C. & Brodlie, M. Airway surface liquid homeostasis in cystic fibrosis: pathophysiology and therapeutic targets. *Thorax* **71**, 284–287 (2016).

43. Floquet, C., Hatin, I., Rousset, J. P. & Bidou, L. Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. *PLoS Genet.* **8**, e1002608 (2012).

44. Iwano, S. et al. Single-cell bioluminescence imaging of deep tissue in freely moving animals. *Science* **359**, 935–939 (2018).

45. de Alwis, R. et al. A single dose of self-transcribing and replicating RNA-based SARS-CoV-2 vaccine produces protective adaptive immunity in mice. *Mol. Ther.* **29**1983 (2021).

46. Sternberg, N., Sauer, B., Hoess, R. & Abremski, K. Bacteriophage P1 cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. *J. Mol. Biol.* **187**, 197–212 (1986).

47. Lin, Q., Jo, D., Gebre-Amlak, K. D. & Ruley, H. E. Enhanced cell-permeant Cre protein for site-specific recombination in cultured cells. *BMC Biotechnol.* **4**, 25 (2004).

48. Schultz, D. et al. 16HBE cell lipid mediator responses to mono and co-infections with respiratory pathogens. *Metabolites* **10**, 113 (2020).

49. Valley, H. C. et al. Isogenic cell models of cystic fibrosis-causing variants in natively expressing pulmonary epithelial cells. *J. Cyst. Fibros.* **18**, 476–483 (2019).

50. Andersen, G. R. et al. Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1Bα. *Mol. Cell* **6**, 1261–1266 (2000).

51. Schrader, J. M., Chapman, S. J. & Uhlenbeck, O. C. Understanding the sequence specificity of tRNA binding to elongation factor Tu using tRNA mutagenesis. *J. Mol. Biol.* **386**, 1255–1264 (2009).

52. Yarus, M., Cline, S. W., Wier, P., Breeden, L. & Thompson, R. C. Actions of the anticodon arm in translation on the phenotypes of RNA mutants. *J. Mol. Biol.* **192**, 235–255 (1986).

53. Giege, R., Sissler, M. & Florentz, C. Universal rules and idiosyncratic features in tRNA identity. *Nucl. Acids Res.* **26**, 5017–5035 (1998).

54. Kao, C., Rudisser, S. & Zheng, M. A simple and efficient method to transcribe RNAs with reduced 3′ heterogeneity. *Methods* **23**, 201–205 (2001).

55. Jurk, M. et al. Human TLR7 or TLR8 independently confer responsiveness to the antiviral compound R-848. *Nat. Immunol.* **3**, 499 (2002).

56. Rajappan, K. et al. Property-driven design and development of lipids for efficient delivery of siRNA. *J. Biol. Chem.* **63**, 12992–13012 (2020).

57. Zomer, A. et al. In vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior. *Cell* **161**, 1046–1057 (2015).

58. Aksit, M. A. et al. Decreased mRNA and protein stability of W1282X limits response to modulator therapy. *J. Cyst. Fibros.* **18**, 606–613 (2019).

59. Sanderlin, E. J. et al. *CFTR* mRNAs with nonsense codons are degraded by the SMG6-mediated endonucleolytic decay pathway. *Nat. Commun.* **13**, 2344 (2022).

60. Amirbeigiarab, S. et al. Invariable stoichiometry of ribosomal proteins in mouse brain tissues with aging. *Proc. Natl Acad. Sci. USA* **116**, 22567–22572 (2019).

61. Kirchner, S. et al. Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol.* **15**, e2000779 (2017).

62. Polte, C. et al. Assessing cell-specific effects of genetic variations using tRNA microarrays. *BMC Genomics* **20**, 549 (2019).

63. Durmowicz, A. G., Lim, R., Rogers, H., Rosebraugh, C. J. & Chowdhury, B. A. The U.S. Food and Drug Administration's experience with ivacaftor in cystic fibrosis. Establishing efficacy using in vitro data in lieu of a clinical trial. *Ann. Am. Thorac. Soc.* **15**, 1–2 (2018).

64. Han, S. T. et al. Residual function of cystic fibrosis mutants predicts response to small molecule CFTR modulators. *JCI Insight* **3**, e121159 (2018).

65. Birket, S. E. et al. Combination therapy with cystic fibrosis transmembrane conductance regulator modulators augment the airway functional microanatomy. *Am. J. Physiol.* **310**, L928–939 (2016).

66. Birket, S. E. et al. A functional anatomic defect of the cystic fibrosis airway. *Am. J. Respir. Crit. Care Med.* **190**, 421–432 (2014).

67. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–840 (2010).

68. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

69. Bartholomaus, A. & Ignatova, Z. Codon resolution analysis of ribosome profiling data. *Methods Mol. Biol.* **2252**, 251–268 (2021).

Acknowledgements

The authors thank J. van Rheenen for the gift of pLV-CMV-LoxP-DsRed-LoxP-eGFP; R. Trelles for the Cre-*loxP* HEK293T reporter cell line; P. Hartman for construction of some reporter mRNA constructs; H. Valley and M. Mense. for the 16HBEge and primary patient-derived hNE cells; and M. Sablad and T. Dam for technical guidance and assistance regarding mouse experiments. This research was supported by grants from the Cystic Fibrosis Foundation (IGNATO2010 to Z.I. and ROWE19R0 to S.M.R.), NIH (1R01HL136414-01 and 1R01HL136414-05 to Z.I. and E.J.S., and P30DK072482 and R35HL135816 to S.M.R.), the German Cystic Fibrosis Foundation muko e.V (2105 to S.A.), Hamburg Innovation C4T projects (C4T635 to Z.I. and S.A.), and a scholarship from ANII Uruguay (POS-EXT-2020-1-164944 to M.D.). Illustrations were created with BioRender.com.

Funding

Open access funding provided by Universität Hamburg.

Author information

Authors and Affiliations

1. Institute of Biochemistry and Molecular Biology, University of Hamburg, Hamburg, Germany

   Suki Albers, Nikhil Bharti, Marcos Davyt, Leonardo Santos, Miguel Angel Delgado-Toscano & Zoya Ignatova

2. Arcturus Therapeutics, San Diego, CA, USA

   Elizabeth C. Allen, Carlos G. Perez-Garcia, Rajesh Mukthavaram, Brandon Molina, Kristen Kuakini, Maher Alayyoubi, Kyoung-Joo Jenny Park, Grishma Acharya, Jose A.

Gonzalez, Amit Sagi, Kiyoshi Tachikawa, Priya Karmali, Daiki Matsuda & Pad Chivukula

3. Department of Pediatrics, School of Medicine, Emory University, Atlanta, GA, USA

   Disha Joshi, Candela Manfredi, Jeong S. Hong & Eric J. Sorscher

4. Children's Healthcare of Atlanta, Atlanta, GA, USA

   Disha Joshi, Candela Manfredi, Jeong S. Hong & Eric J. Sorscher

5. Pulmonary, Allergy, and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

   Susan E. Birket & Steven M. Rowe

6. Wellman Center for Photomedicine, Massachusetts General Hospital, Boston, MA, USA

   Guillermo J. Tearney

7. Harvard-MIT Health Sciences and Technology, MA, Cambridge, USA

   Guillermo J. Tearney

Contributions

S.A., K.T., P.C., D.M. and Z.I. conceptualized the work and strategy. S.A., E.C.A., N.B., M.D., D.J., C.G.P.-G., M.A.D.-T., B.M. and D.M. performed the experiments and analysed the data. L.S. and Z.I. analysed the deep-sequencing data. E.C.A., K.K., K.J.-J.P., J.A.G. and M.A. performed the synthesis of tRNA and mRNA for in vivo formulations and G.A. performed the administrations in mouse. R.M., A.S. and P.K. made the various LUNAR–mRNA and LUNAR–tRNA formulations. S.E.B., G.J.T. and S.M.R. performed the μOCT measurements. J.S.H. produced the FRT cell line, stably expressing *CFTR*^{R553X} and with C.M. and D.J. performed the channel activity measurements in primary cells. S.A., E.C.A., E.J.S., D.M., K.T., P.C. and Z.I. planned the experiments. S.A., E.C.A., N.B., M.D., D.J.,

C.G.P.-G., J.S.H., C.M., E.J.S., G.A., D.M. and Z.I. analysed the experiments and interpreted the data. S.A. and Z.I. supported by N.B., M.D., L.S., D.M. and E.J.S. wrote the manuscript. All authors supported the review of the final version of the manuscript.

Corresponding authors

Correspondence to Eric J. Sorscher, Pad Chivukula or Zoya Ignatova.

Ethics declarations

Competing interests

Z.I., S.A., N.B. and M.D. are inventors on patents related to tRNA designs for PTC correction. Z.I. is also a scientific advisor for Tevard Biosciences. S.M.R. and G.J.T. are named on an unlicensed patent on the use of OCT for airway surface liquid measurements. S.M.R. is named on an unrelated patent on translational readthrough pharmacotherapy. E.J.S. is a non-voting board member of the Cystic Fibrosis Foundation. The LUNAR technology is proprietary to Arcturus Therapeutics. E.C.A., C.G.P.-G., R.M., B.M., K.K., M.A., K.J.-J.P., G.A., J.A.G., A.S., K.T., P.K., D.M. and P.C. are employees and have securities from Arcturus Therapeutics. The other authors declare no competing interests.

Peer review

Peer review information

Nature thanks Christopher Trotta and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 tRNA designs to suppress PTCs at Gly codons.

a, Cloverleaf structure (left) of the human tRNA^{Gly}UCC which is represented by two isodecoders that share the same anticodon to decode GGA codon, but differ by a single nucleotide (red) in their sequence. The anticodon of both tRNA^{Gly} isodecoders were mutated to pair to the UGA PTC yielding tG1 and tG2 (Supplementary Table 1); nucleotide substitutions in TΨC stem yielding the Ti variants (right) are highlighted in red. **b**, Suppression efficiency of tG variants at UGA PTC monitored with R208X-FLuc in Hep3B cells and normalized to wild-type FLuc expression. In vitro transcribed tG variants were co-transfected with plasmid-borne $FLuc^{R208C}$ reporter construct (schematic Fig. 1b). Anticodon-edited tG1 and tG2, black; TΨC-stem edited variants of tG1 and tG2, red; mismatch tRNA that does not pair to UGA PTC, dark blue. Data are means ± s.e.m. (n = 4 biologically independent experiments).

Source Data

Extended Data Fig. 2 In vitro transcription generates high quality functional tRNAs with functionally homogenous 3′ ends in Hepa1-6 cells.

a, **b**, Purified in vitro transcribed tRNAs are homogenous as monitored by denaturing polyacrylamide gel (**a**) or capillary electrophoresis (**b**) as exemplified by tS and tSA1T5. Left lane on the polyacrylamide gel, RNA ladder (Low Range ssRNA ladder, NEB). LM, lower marker. Analysis performed once with a randomly chosen batch. **c**, Readthrough efficacy of tSA1T5 transcribed with conventional primer (closed circles) is comparable to that of the tSA1T5 produced using 3′-modified (2′-OMe) primer (open circles) which enables precise termination of in vitro tRNA transcription. Data are means ± e.e.m. (n = 3 biologically independent experiments). For gel source data to panel a see Supplementary Fig. 2.

Source Data

Extended Data Fig. 3 tSA1T5 incorporates only Ser in place of the UGA stop codon.

a, b, Hepa 1-6 cells co-transfected with in vitro transcribed tSA1T5 and *FLuc*^{R208X}-*RLuc* mRNA (**a**) or WT *FLuc*^{R208}-*RLuc* mRNA (**b**) and analyzed by 2D-nano PRM LC-MS/MS. Extracted Ion Chromatograms (EICs) (**a,b** upper panels) and MS/MS spectra (**a,b** lower panels). Peptides with m/z 685.8089 and 642.2929 correspond to the double-charged Luc-derived peptides. Peptide **S**YNLTSQDEDAK is identified in R208X-FLuc expressing cells transfected with tSA1T5, indicating the incorporation of Ser at the UGA stop codon. Sequence parts of tSA1T5 originate from tRNA^{Thr}GCU, yet they do not compromise the fidelity of seryl-tRNA-synthetase, as no threonine incorporation was detected (Supplementary Table 1). No tryptophan or selenocysteine was detected either despite the intrinsic tendency of tRNA^{Trp} or tRNA^{Sec} to pair to UGA codon. Wild-type FLuc-derived peptides from the FLuc-R208-RLuc reporter without PTC were cleaved directly after the Arg208 residue generating the peptide YNLTSQDEDAK. The experiment serves as a quality check and hence, was performed as a single experiment.

Extended Data Fig. 4 Dose-dependent suppression activity of sup-tRNA and cell viability.

a, Suppression activity monitored by co-transfecting different concentrations of in vitro transcribed tS (black) or tSA1T5 (orange) suppressor tRNAs . PTC suppression activity is represented as a ratio of the normalized PTC-FLuc reporter activity from *FLuc*-^{R208X}-*RLuc* mRNA to that of the reporter without PTC (FLuc(R208S)-RLuc). Data are means ± s.e.m. (n = 3 biologically independent experiments). **b**, Cell viability assay 24 h post transfection of CFBE41o^{-} cells transfected with different tS (black) and tSA1T5 (orange) concentrations. The signal was normalized to mock-transfected cells and the signal with the lowest tRNA concentration was set as 100%. Data are means ± s.e.m (n = 3 biologically independent experiments). **c**, Schematic of screening by co-transfecting of in vitro

transcribed (IVT) sup-tRNA and PTC-*FLuc*-*RLuc* reporter mRNA. FLuc signal was normalized on the RLuc, both of which encoded within the same reporter but expressed from independent promoters. *RLuc* uses cap-independent IRES . **d**, tSA1T5 targets PTCs with different stop codon identities tested in murine Hepa 1-6 cells in the setup introduced in panel c. PTC suppression activity was calculated using the ratio of the normalized PTC-FLuc reporter activity from FLuc(R208X)-RLuc mRNA to that of a reporter without PTC (FLuc-R208S-RLuc), and represented as % readthrough efficiency at the PTC. Data are means ± s.e.m. (n = 3 biologically independent experiments). $^{**}P < 0.01$, $^{***}P < 0.001$ (one-sided *t* test). **e**, Comparison of PTC suppression efficiency by co-transfecting tSA1T5 (orange) with the PTC (UGA) reporter as either mRNA (**c**), or DNA (schematic, Fig. 1b). mis (dark blue) designates mismatch tRNA that does not pair to any PTC. Data are means ± s.e.m. (for mRNA – n = 3, for DNA – n = 4, for mis combined n = 7 biologically independent experiments).

Source Data

Extended Data Fig. 5 Stability of the administered mRNA and sup-tRNA.

a, The amount of the control $aLuc^{R208S}$ mRNA inferred by aLuc activity in mice decreased by two orders of magnitude within 24 h. Quantification of aLuc activity expressed from $aLuc^{R208S}$ mRNA post i.v. injection from the IVIS images of the Balb/C mice (Fig. 2b) at 6 h (open bars) or 24 h (filled bars). $aLuc^{R208S}$ mRNA was co-encapsulated with tS (groups 7, 8) or tSA1T5 (groups 9, 10) at 0.6 mg/kg sup-tRNA (group 7, 9) or 1.2 mg/kg sup-tRNA (group 8, 10). Data are means ± s.e.m. (n = 3 mice in each group). **b**, Representative block from the tRNA-tailored microarray to detect sup-tRNA with high sensitivity. Full-length tDNA probes complementary to tSA2T5 were used to detect sup-tRNA (white squares). Each array consists of 12 such blocks, each of which contains a duplicate of three distinct probes recognizing all $tRNA^{Ser}$ isoacceptors, i.e. $tRNA^{Ser}$(wGA), a combined probe for $tRNA^{Ser}$AGA and $tRNA^{Ser}$UGA (recognizing UCU, UCC and UCA codons; both probes at the bottom right), $tRNA^{Ser}$GCU (pairing to AGU and AGC codons; probes - upper right), and $tRNA^{Ser}$CGA (reading codon UCG, bottom left flanking the spike-ins). Blue squares, spike-ins or non-human

tRNAs, used for signal normalization. tSA2T5 share some sequence similarity with $tRNA^{Ser}$AGA and $tRNA^{Ser}$UGA whose backbone was the starting sequence for tSA2T5 (Supplementary Table 1). In the control group, untreated with sup-tRNA, none of the natural $tRNA^{Ser}$ isoacceptors (purple squares) hybridize to tSA2T5 probe suggesting high specific of detection. **c**, sup-tRNA stability in gene-edited 16HBEge cells (16HBEge$^{R1162X/-}$) monitored with tRNA microarrays represented as box plot (n = 2 independent replicates) and normalized to the mean of the signal at first time point (5 h) post transfection which was set as 100 % (horizontal line). The center line indicates the median, box ends are 10%-90%, and the whiskers – the range of the remaining data without exclusion of outliers.

Source Data

Extended Data Fig. 6 Cre-lox system is suitable to monitor PTC suppression.

a, Schematic of the Cre-lox system. eGFP will be only expressed if full-length Cre recombinase is expressed, which through subsequent recombination removes DsRed gene from the original cassette. Each fluorescent protein is color coded and captures the corresponding fluorescence pattern on the images. Arrow and triangle designate the CMV and *loxP* promoters, respectively. **b, c** Time course of DsRed and eGFP expression HEK293T cells stably expressing the floxed-DsRed-floxed-eGFP cassette and transfected with *Cre* mRNA and tRNA variants. Integrated fluorescence (**b**) and representative microscopy images at day 4 post transfection (**c**). A PTC-*Cre* mRNA with two UGA PTCs (S69X-S82X; blue) was used for co-transfection with tS variants designed to suppress UGA, UAA or UAG codons, respectively. Wild-type *Cre* mRNA without PTC (black) and mock (m) transfected cells (grey) served as positive and negative control of Cre-*lox* recombination, respectively. Notably, we observed elevation of eGFP signals along with simultaneous reduction of DsRed only in cells co-transfected with PTC-*Cre* mRNAs and PTC-matching tRNA (tS::UGA) suggesting efficient PTC correction and expression of full-length Cre recombinase. The eGFP signal decreased at day 4, likely because of degradation and/or decreased protein expression in

general, since the DsRed signal also decreased. Data are means ± s.e.m. (n = 3 biologically independent experiments). Scale bar, 200 µM.

Source Data

Extended Data Fig. 7 In vivo PTC suppression in lung by LUNAR-formulated tSA1T5 in a transgenic Flox-tdTomato mouse model.

Brightfield immunohistochemical images for tdTomato expression in lungs from mice (n = 4 mice in each group) intratracheally dosed with LUNAR2-LNPs that were formulated with WT-*Cre* mRNA (no PTC), PTC-*Cre* mRNA with UGA at S69X-S82X with or without tSA1T5 with an anticodon pairing to UGA or a mismatch tRNA (mis). Control mice with PTC-*Cre* mRNA administered only were used to establish any background tdTomato that might be present as reported by the transgenic line producer (Jackson Laboratories, USA). Mouse lung treated with PBS served as a negative control. Without a suppressor tRNA, PTC-$Cre^{S69X\text{-}S82X}$ mRNA did not result in tdTomato expression and is indistinguishable from the negative control (PBS). tdTomato expression is detectable as dark brown spots and seen only in the WT-*Cre* mRNA or following suppression with tSA1T5. Representative image from each mouse of each group is included. Square in the whole image (bottom right) points location of each panel. Scale bar, 250 µM. Successful PTC suppression is indicated by the higher numbers of cells expressing tdTomato (red arrow on the magnified images, scale bar, 100 µM). Very low background of tdTomato expression from mice dosed with PTC-Cre$^{S69X\text{-}S82X}$ mRNA and non-matching mismatch tRNA (mis) was observed (red arrow on the magnified images, scale bar, 100 µM), likely due to a low Cre-independent tdTomato background reported for these transgenic mice. Images designated with * in the right corner are the ones shown in Fig. 2d–g.

Extended Data Fig. 8 Suppressor tRNAs do not induce readthrough at native UGA codons in mice and cell culture.

a-b, Marginal but comparable readthrough at all three natural stop codons (including the unspecific ones UAA and UAG not targeted by the sup-

tRNA) detected in the cumulative coverage plots (upper panels), the ribosome readthrough score (RRTS, lower panels) and Venn diagrams of the top 10% transcripts with the highest RRTS (on the right) in lung tissue of mice treated by i.t. microsprayer instillation with nanoparticulated tSA1T5 (**a**), liver tissue from mice treated by i.v. with nanoparticulated tSA1T5 (**b**). The center line indicates the median, the box ends – the first and third quartiles, and the whiskers – the range of the remaining data without exclusion of outliers. (**b**). Insets, box-plots as in the main plot without the outliers to visualize the median values. Control, mice treated with PBS; R#1 and R#2 indicate two mice used as independent replicates. The RRTS is a comparison between tRNA-treated samples compared to the corresponding tissue of a control mice (**a,b**). We noted that all detected transcript with a high RRTS (Venn diagrams) are within 10%ile of genes with highest expression level, suggesting a high proportion of false positives. **c,d**, Suppressor tRNA do not enhance the readthrough at internal UGA stop codons. RRTS in lung tissue of mice treated by i.t. with nanoparticulated tSA1T5 (**c**) and liver tissue from mice treated by i.v. with nanoparticulated tSA1T5 (**d**). The RRTS is a comparison between tRNA-treated samples compared to the corresponding tissue of a control mice (**c,d**). Note that different number of genes with internal UGA stop codons were detected as expressed in different tissues (i.e. 11 internal stop codons in mouse liver, 2 in lung tissue).

Source Data

Extended Data Fig. 9 Efficiency of tS and tR variants in suppressing PTCs in CFTR-expressing immortalized cell lines.

a,b, Suppression activity in full-length CFTR was monitored by automated immunoblotting (JESS system) using monoclonal CFTR-NBD2 antibody (1:100 dilution, #596) in CFBE41o^{-} cells co-transfected with plasmid-encoded $CFTR^{S466X}$, $CFTR^{R553X}$, $CFTR^{R1162X}$ and the corresponding in vitro transcribed tRNA variant, using lipofectamine as transfection reagent. The immunofluorescence signal of the mature fully glycosylated CFTR (band C) was normalized against signal from the total protein staining (blue). mis, PTC-CFTR variants co-transfected with mismatch tRNA which does not pair to the PTC. Lower molecular weight bands are likely CFTR degradation

or incompletely synthesized products. **b, c**, Representative Ussing chamber tracings obtained from FRT^{R553X} (**b**) and FRT^{R1162X} (**c**) cells transfected with tRT5 (red) or mismatch tRNA (blue) using lipofectamine. Forskolin, CFTR activator; VX-770, potentiator, Inh-172, CFTR inhibitor. Statistically insignificant (minimally negative) VX-770 deflections were noted in some tracings after peak CFTR stimulation by forskolin. Gel source data to panel a included in Supplementary Fig.1.

Source Data

Extended Data Fig. 10 Restoration of CFTR expression and activity by tRT5 or by co-treatment with NMD inhibitor in patient-derived hNE (R1162X/R1162X) cells.

a, Representative immunoblots (automated immunoblotting, JESS system) using monoclonal CFTR-NBD2 antibody (1:100 dilution, #596) in 16HBEge$^{R1162X/-}$ and hNE$^{R1162X/R1162X}$ transfected with sup-tRNA alone using lipofectamine or treated with 5µM NMD14. PTC124 (ataluren). **b**, Wild-type (WT) CFTR expression (band C normalized to total protein) in hNE derived from healthy individuals with and without treatment with NMD_i, (NMD14 (5 µM)) for 24h. Data are means ± s.e.m. (n = 3 biologically independent replicates). Statistics, one-sided t test. NMD alone decreases the WT-CFTR expression most likely owning to alterations of hNE viability following NMD treatment. **c**, Representative short-circuit current traces obtained from hNE$^{R1162X/R1162X}$ cells transfected with tRT5 (red) or mismatch tRNA (blue). Forskolin, CFTR activator; VX-770, potentiator, Inh-172, CFTR inhibitor. Gel source data to panel a included in Supplementary Fig. 1.

Source Data

Supplementary information

Supplementary Information

This file contains Supplementary Figs. 1a–c and 2 and Supplementary Tables 1–4.

Reporting Summary

Peer Review File

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Figs. 1, 2, 4, 5, 6, 8–10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Albers, S., Allen, E.C., Bharti, N. *et al.* Engineered tRNAs suppress nonsense mutations in cells and in vivo. *Nature* **618**, 842–848 (2023). https://doi.org/10.1038/s41586-023-06133-1

- Received: 31 July 2021
- Accepted: 25 April 2023
- Published: 31 May 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06133-1

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06133-1

- Article
- Open Access
- Published: 07 June 2023

A cytosolic surveillance mechanism activates the mitochondrial UPR

- F. X. Reymond Sutandy,
- Ines Gößner,
- Georg Tascher &
- …
- Christian Münch

Show authors

Nature **volume 618**, pages 849–854 (2023)

- 9846 Accesses
- 116 Altmetric
- Metrics details

Abstract

The mitochondrial unfolded protein response (UPR^{mt}) is essential to safeguard mitochondria from proteotoxic damage by activating a dedicated transcriptional response in the nucleus to restore proteostasis[1,2]. Yet, it remains unclear how the information on mitochondria misfolding stress (MMS) is signalled to the nucleus as part of the human UPR^{mt} (refs. [3,4]). Here, we show that UPR^{mt} signalling is driven by the release of two individual signals in the cytosol—mitochondrial reactive oxygen species (mtROS) and accumulation of mitochondrial protein precursors in the

cytosol (c-mtProt). Combining proteomics and genetic approaches, we identified that MMS causes the release of mtROS into the cytosol. In parallel, MMS leads to mitochondrial protein import defects causing c-mtProt accumulation. Both signals integrate to activate the UPR^{mt}; released mtROS oxidize the cytosolic HSP40 protein DNAJA1, which leads to enhanced recruitment of cytosolic HSP70 to c-mtProt. Consequently, HSP70 releases HSF1, which translocates to the nucleus and activates transcription of UPR^{mt} genes. Together, we identify a highly controlled cytosolic surveillance mechanism that integrates independent mitochondrial stress signals to initiate the UPR^{mt}. These observations reveal a link between mitochondrial and cytosolic proteostasis and provide molecular insight into UPR^{mt} signalling in human cells.

Main

Maintenance of mitochondrial protein homoeostasis is crucial for mitochondrial function. Upon proteotoxic stress, mitochondria activate the mitochondrial unfolded protein response (UPR^{mt}), a nuclear transcriptional response that induces mitochondrial chaperones, such as *HSPD1*, *HSPE1* and *HSPA9*, and proteases, including *LONP1*, to re-establish homoeostasis in mitochondria[2,3,5]. The molecular events underlying the retrograde mitochondria–nucleus communication to induce the UPR^{mt} in humans remain unclear. The integrated stress response (ISR) has been shown to contribute to the cellular rearrangements observed during the UPR^{mt} and during mitochondrial stress responses in general[2,6,7,8]. However, recent findings indicated that the mitochondrial stress response/ISR and the UPR^{mt} are two independent processes that are part of a more complex stress response[1,9,10,11].

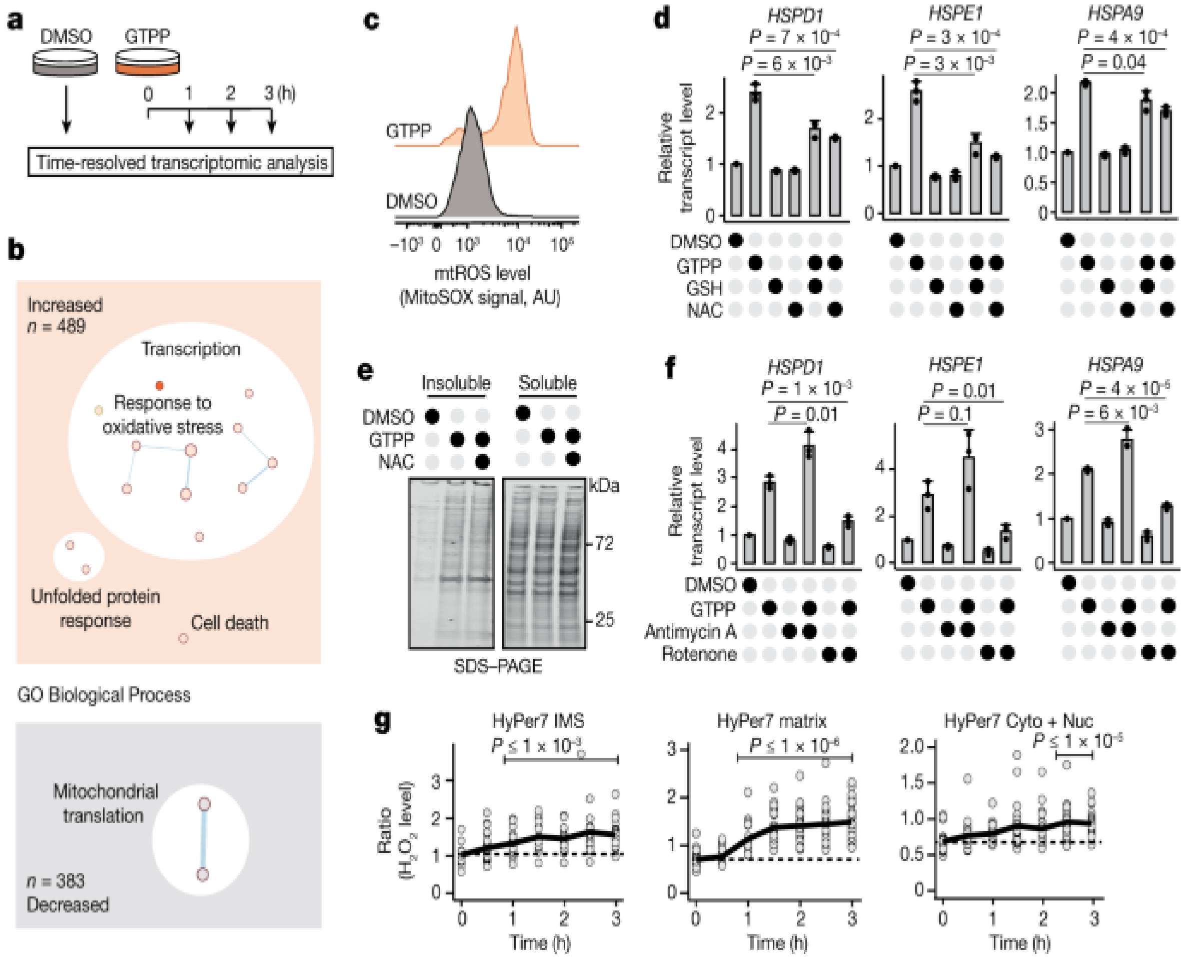
To study the role of the ISR in UPR^{mt} signalling, we monitored the early responses to mitochondrial misfolding stress (MMS). Treatment with the mitochondrial HSP90 inhibitor gamitrinib-triphenylphosphonium (GTPP)[1] causes MMS and significantly induced the UPR^{mt} genes within 2–3 h but not general mitochondrial genes (Extended Data Fig. 1a–c). The primary ISR effector *ATF4* was induced before the UPR^{mt}, while *CHOP* (a direct target of ATF4) induction showed a similar profile to UPR^{mt} genes

(Extended Data Fig. 1a,d). However, knockout (KO) of both main ISR effectors did not reduce the UPR^{mt}, suggesting that the ISR–ATF4 axis was not required for UPR^{mt} induction (Extended Data Fig. 1e–h).

Mitochondrial reactive oxygen species are required for UPR^{mt} activation

To identify the molecular signatures that signal the UPR^{mt}, we carried out time-resolved transcriptomic analyses of cells within 3 h of GTPP treatment (Fig. 1a). Principal component analysis revealed that cells treated with GTPP showed distinct transcriptomic patterns over time (Extended Data Fig. 2a), indicating a dynamic transcriptional response during UPR^{mt} activation, with 489 and 383 transcripts gradually increased and decreased, respectively (Fig. 1b, Extended Data Fig. 2b–d and Supplementary Table 1). Genes prominently enriched during UPR^{mt} activation included ‘response to oxidative stress’ (Fig. 1b and Extended Data Fig. 2c,e), suggesting that reactive oxygen species (ROS) may contribute to UPR^{mt} signalling. In line with this hypothesis, induction of MMS caused increased mitochondrial reactive oxygen species (mtROS; $O_2^{\bullet-}$) levels (Fig. 1c and Extended Data Fig. 2f). While high levels of ROS can be detrimental, mitochondria often use ROS to communicate with different organelles[12]. To test whether ROS are necessary for UPR^{mt} activation, we carried out cotreatments with the antioxidants *N*-acetylcysteine (NAC) and reduced glutathione (GSH) and the superoxide dismutase mimetic MnTBAP. Strikingly, all three antioxidants inhibited UPR^{mt} induction (Fig. 1d and Extended Data Fig. 2g–l) without affecting mitochondrial protein aggregate formation (Fig. 1e). The opposite experimental paradigm, cotreatment with the complex III inhibitor antimycin A to increase mtROS, enhanced UPR^{mt} activation (Fig. 1f and Extended Data Fig. 2m). These findings show that mitochondria employ mtROS as an essential signal to activate and scale the UPR^{mt}. However, increasing mtROS levels alone was not sufficient to activate the UPR^{mt} (Fig. 1f and Extended Data Fig. 2m), indicating that additional factors are required for UPR^{mt} activation.

Fig. 1: ROS are produced and required to activate the UPR^{mt} upon MMS.



a, Scheme of the experimental design for time-resolved transcriptomics (RNA sequencing) upon GTPP treatment. Dimethyl sulfoxide (DMSO) was used as negative control. **b**, Enrichment maps of GO Biological Process (BP) terms from the transcriptomic analysis. Circles represent GO terms with a false discovery rate of less than 0.1. GO BP 'transcriptional response to oxidative stress' (GO:0036091) is marked with a dark orange circle. The total numbers of transcripts belonging to increased and decreased groups are represented as *n*. **c**, Representative FACS measurement of mtROS levels with MitoSOX upon GTPP treatment. AU, arbitrary unit. **d**, Bar plots showing the mean of relative transcript levels of UPR^{mt} genes of GTPP-treated cells upon cotreatments with the antioxidants NAC and GSH measured with qPCR ($n = 3$ biological replicates). **e**, Gel image of mitochondrial insoluble and soluble fractions upon different treatments. Insoluble fractions are a measure for aggregate formation. **f**, Bar plots

showing the mean of relative transcript levels of UPR^{mt} genes of GTPP-treated cells upon cotreatments with mtROS inducers antimycin A and rotenone measured with qPCR ($n = 3$ biological replicates). **g**, Microscopy-based measurement of H_2O_2 levels with HyPer7 reporters targeted to the IMS and the matrix and untargeted (cytosol + nucleus (Cyto + Nuc)) upon GTPP treatment. Each point represents an individual cell measurement. Lines represent the mean H_2O_2 levels across different time points ($n = 18$ cells) during 3 h of measurement from five biological replicates. All *P* values are calculated with a two-tailed unpaired Student's *t* test and indicated. All error bars represent mean ± s.d.

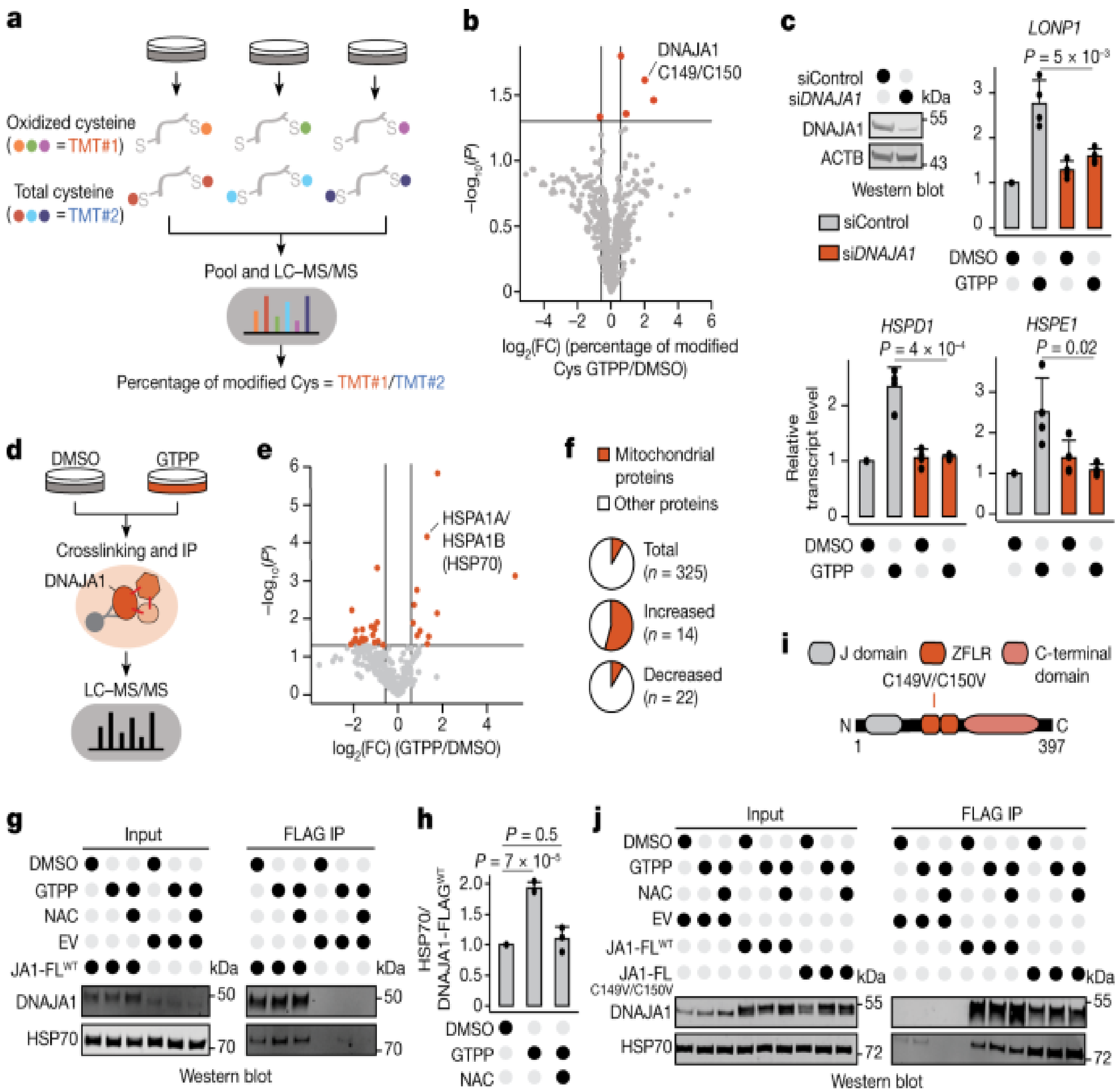
Source data

Next, we investigated the source of mtROS to understand how mtROS mediate UPR^{mt} activation. Intriguingly, we found that elevating mtROS production by cotreatment with the complex I inhibitor rotenone had the opposite effect to antimycin A (Fig. 1f and Extended Data Fig. 2m), indicating site specificity of mtROS production to signal the UPR^{mt}. To monitor compartment-specific mtROS production, we used the ultrasensitive H_2O_2 probe HyPer7 (ref. [13]) targeted to different mitochondrial compartments, as H_2O_2 is the most common type of ROS used in intracellular signalling[14] (Extended Data Fig. 3a–f). H_2O_2 levels increased significantly in the intermembrane space (IMS) and the matrix within 1 h of GTPP treatment (Fig. 1g). At later time points, H_2O_2 levels also significantly increased in the cytosol (Fig. 1g and Extended Data Fig. 3c–f). These findings support a model in which mtROS diffuse into the cytosol and signal the UPR^{mt}. Consistently, blocking ROS transport between mitochondria and the cytosol with 4,4′diisothiocyanatostilbene-2,2′-disulfonate (DIDS), an inhibitor of the outer membrane pore VDAC1, abolished UPR^{mt} activation (Extended Data Fig. 3g–i). Together, our findings show that mtROS accumulation and diffusion into the cytosol are essential for UPR^{mt} signalling.

DNAJA1 oxidation regulates the UPR^{mt}

We considered that mtROS produced during MMS oxidize a cytosolic target to mediate UPR^{mt} signalling. To identify proteins oxidized upon UPR^{mt} activation, we carried out unbiased, multiplexed redox proteomics to identify cytosolic proteins that are reversibly cysteine oxidized upon GTPP treatment (Fig. 2a and Extended Data Fig. 4a). We monitored oxidation changes within 3 h of MMS induction to identify changes involved in early UPR^{mt} signalling. Four cysteine residues showed increased oxidation upon GTPP treatment (Fig. 2b, Extended Data Fig. 4b and Supplementary Table 2). Intriguingly, one of the proteins identified to be oxidized during UPR^{mt} activation was the cytosolic HSP40 (DNAJA1), a cochaperone of cytosolic HSP70 (ref. [15]). We found increased DNAJA1 oxidation at cysteines 149 and 150 upon GTPP treatment (Fig. 2b). Oxidation of DNAJA1 and its homologues has been shown to influence its activity by regulating its zinc finger-like regions (ZFLRs)[16,17]. This renders DNAJA1 sensitive to redox changes, presenting a potential target for the redox signalling of the UPR^{mt}. Indeed, depletion of DNAJA1, but not other HSP40 members, prevented UPR^{mt} activation (Fig. 2c and Extended Data Fig. 4c–e), showing that DNAJA1 is essential for UPR^{mt} signalling.

Fig. 2: DNAJA1 mediates the signalling to activate the UPR^{mt}.



a, Scheme of the redox proteomics for cells treated with DMSO or GTPP (in biological triplicates). Oxidized and total cysteine side chains were labelled with different iodoTMTs, pooled into one six-plex sample and measured by LC–MS/MS. **b**, Volcano plot of the redox proteome showing changes in cysteine side-chain oxidation upon GTPP treatment. Dark orange points represent significantly changed cysteine side chains ($P \leq 0.05$, fold change (FC) > 1.5). Statistical analysis was performed with a two-tailed unpaired Student's *t* test. **c**, Bar plots showing the mean of relative transcript levels of UPR^{mt} genes of GTPP-treated cells upon knockdown of *DNAJA1* measured with qPCR ($n = 4$ biological replicates). The knockdown efficiency is shown in the western blot image (left). **d**, Scheme of the experimental steps for

DNAJA1 quantitative interaction proteomics. IP, immunoprecipitation. **e**, Volcano plot of the DNAJA1 interaction proteomics upon GTPP treatment. Dark orange points represent significantly changed interactions ($P \leq 0.05$, FC > 1.5). Statistical analysis was performed with a two-tailed unpaired Student's *t* test. **f**, Pie charts representing the proportions of mitochondrial proteins on different groups of DNAJA1 interacting partners. The total numbers of proteins from individual groups are indicated. **g**,**h**, Representative western blot images of wild-type FLAG-tagged DNAJA1 (JA1-FL^{WT})–HSP70 interactions upon different conditions (**g**) and quantification of three biological replicates (**h**). EV represents the empty vector control. **i**, Schematic representation of DNAJA1 domain composition. The exact positions of DNAJA1 mutations used in the experiments are indicated. **j**, Western blot images of the wild-type, C149V and C150V FLAG-tagged DNAJA1 mutant interaction with HSP70 upon different conditions ($n = 2$ biological replicates). All *P* values are calculated with a two-tailed unpaired Student's *t* test and indicated. All error bars represent mean ± s.d. Gel source data are in Supplementary Figs. 1 and 2.

Source data

To better understand DNAJA1's role in UPR^{mt} signalling, we examined its activity during MMS. Quantitative interaction proteomics showed that DNAJA1 exhibited significantly increased binding to a large number of mitochondrial proteins and the cytosolic HSP70s (HSPA1A and HSPA1B) during GTPP treatment (Fig. 2d–f, Extended Data Fig. 5a,b and Supplementary Table 3). Notably, the DNAJA1–HSP70 interaction was ROS dependent (Fig. 2g,h). DNAJA1 cysteines 149 and 150 are part of the ZFLR (Fig. 2i). Their oxidation during the UPR^{mt} may interfere with zinc ion binding[16]. To assess whether HSP70 recruitment to DNAJA1 during GTPP treatment was mediated by conformational changes of its ZFLR, we introduced C149V and C150V mutations to mimic the effect of oxidation by removal of the cysteines required for the interaction with zinc ions (Fig. 2i). Indeed, mimicking DNAJA1 oxidation increased the DNAJA1–HSP70 interaction, similar to the effect we observed during GTPP treatment (Fig. 2j). These findings suggest that the MMS-induced mtROS lead to oxidation of the DNAJA1 ZFLR to increase HSP70 recruitment.

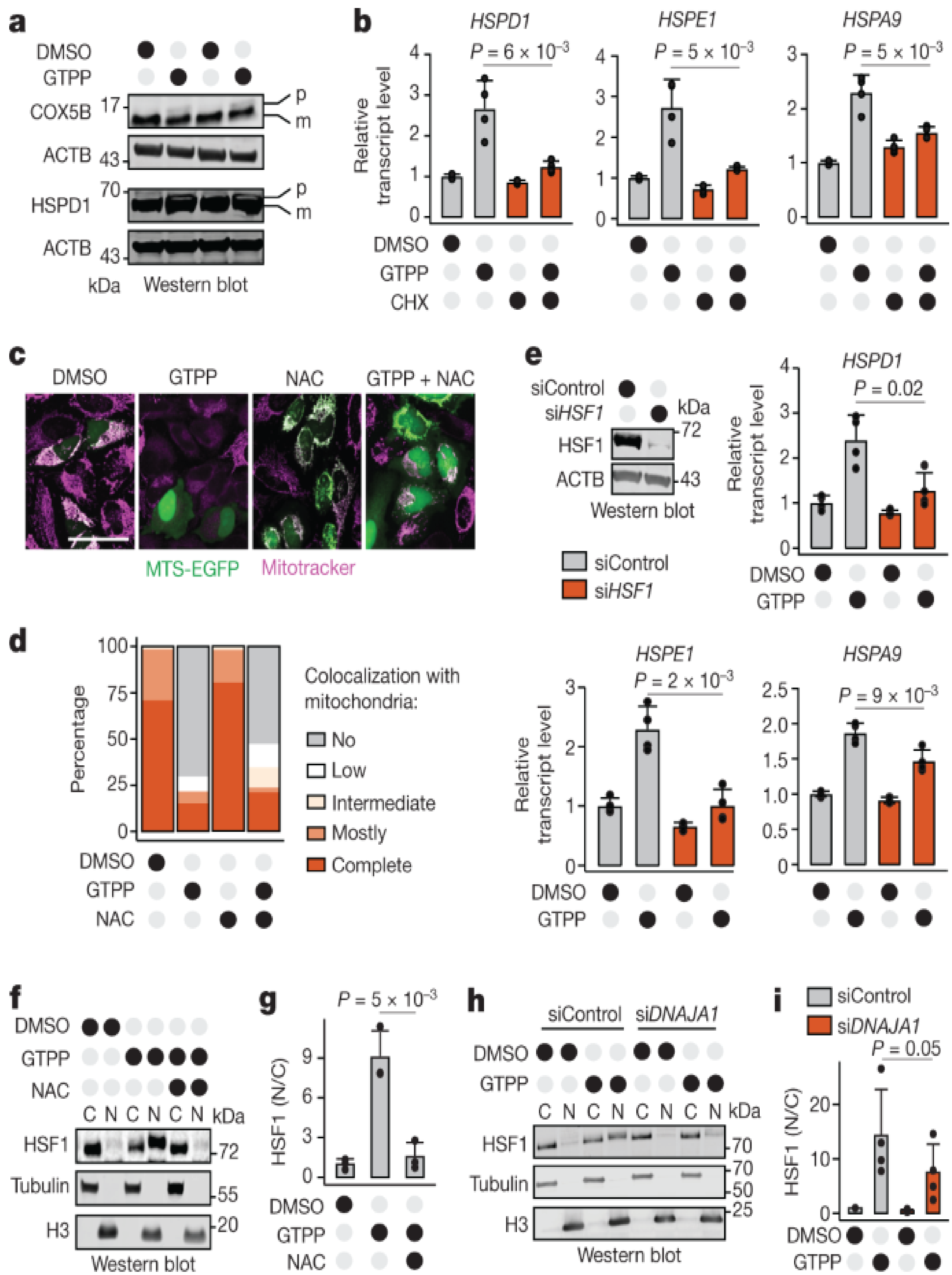
DNAJA1 preselects and delivers specific client proteins to HSP70 (ref. [15]). Thus, the increase in DNAJA1–HSP70 interaction might indicate formation of an active DNAJA1–client complex mediating binding to HSP70. This is consistent with the observed increase in interactions between DNAJA1 and mitochondrial proteins during GTPP treatment (Fig. 2f and Extended Data Fig. 5b). The DNAJA1–HSP70 interaction with mitochondrial proteins occurred in the cytosol, ruling out a potential mislocalization of DNAJA1 to mitochondria during MMS (Extended Data Fig. 5c–e). The presence of mitochondrial proteins in the cytosol was not associated with apoptotic cell death (Extended Data Fig. 5f,g). Together, our results identify DNAJA1 as an integral component of UPR^{mt} signalling.

The UPR^{mt} requires mitochondrial protein precursor accumulation in the cytosol

Next, we evaluated the underlying reasons for the increased interaction of DNAJA1–HSP70 with mitochondrial proteins during UPR^{mt} activation. The majority of mitochondrial proteins are synthesized in the cytosol as precursors that need to be imported into mitochondria, where they are processed into their mature form[18]. In yeast, mitochondrial protein precursors can accumulate in the cytosol during stress and cause activation of cytosolic stress responses that aim at restoring proteostasis[19,20,21]. Whether such mechanisms exist in humans is unclear. We speculated that similar precursors may accumulate in the cytosol and serve as DNAJA1 clients during MMS in humans. Indeed, mitochondrial protein precursors increased during MMS (Fig. 3a). Given that these mitochondrial proteins showed increased interaction with DNAJA1 upon GTPP treatment, we checked whether accumulation of mitochondrial protein precursors in the cytosol (c-mtProt) was required for UPR^{mt} signalling. Preventing c-mtProt accumulation by inhibiting cytosolic protein translation with cycloheximide (CHX) decreased UPR^{mt} activation (Fig. 3b). This effect appeared to be selective for mitochondrial proteins since preventing transcription of mitochondrial genes (but not mitochondrial chaperones) via depletion of the mitochondrial biogenesis factor *NRF1* was sufficient to reproduce these

effects (Extended Data Fig. 6a–c). These observations show that c-mtProt accumulation is a key component of UPR^{mt} signalling.

Fig. 3: ROS and accumulation of c-mtProt activate the DNAJA1–HSF1 axis to induce the UPR^{mt}.



a, Western blot images of mitochondrial proteins in their precursor (p) and mature (m) forms upon GTPP treatment. **b**, Bar plots showing the mean of

relative transcript levels of UPR^{mt} genes of GTPP-treated cells upon cotreatment with CHX measured with qPCR (*n* = 4 biological replicates). **c**, Representative microscopy images of MTS-EGFP (green) localization in comparison with mitochondria (magenta) upon different treatments. Scale bar, 50 µm. **d**, Bar plot depicting mean quantification of microscopy images (*n* = 100 cells). **e**, Bar plots showing the mean of relative transcript levels of UPR^{mt} genes of GTPP-treated cells upon knockdown of *HSF1* measured with qPCR (*n* = 4 biological replicates). The knockdown efficiency is shown by western blotting (upper left). **f**, Representative western blot images of HSF1 in the cytosolic (C) and nuclear (N) fractions of cells under different treatments. **g**, Bar plot depicting the mean of the nuclear-to-cytosolic ratio (N/C) of HSF1 from triplicate images of the western blots. **h**, Representative western blot images of HSF1 in C and N fractions of cells upon knockdown of *DNAJA1*. **i**, Bar plot depicting the mean of the N/C of HSF1 from four replicate images of the western blots. *P* values are calculated with a two-tailed unpaired (**b**,**e**,**g**) or paired (**h**) Student's *t* test and indicated. All error bars represent mean ± s.d. Gel source data are in Supplementary Figs. 1 and 2.

Source data

Our previous work had shown that different stressors compromising mitochondrial import cause c-mtProt accumulation[22]. To check whether inducing MMS leads to mitochondrial import defects, we used an MTS-EGFP reporter[23]. During GTPP treatment, we observed a decreased mitochondrial MTS-EGFP signal and in parallel, an increased signal in the cytosol and nucleus (Fig. 3c,d). These import defects were ROS independent (Fig. 3c,d). Monitoring newly synthesized Halo-tagged mitochondrial proteins resulted in the same observations (Extended Data Fig. 6d–j). Consequently, accumulation of c-mtProt constitutes a second signal of the UPR^{mt}, in addition to mtROS.

mtROS and c-mtProt activate DNAJA1–HSF1

We next addressed which downstream factor might integrate the mtROS and c-mtProt accumulation signals to convey the UPR^{mt} to the nucleus. In yeast, c-mtProt accumulation has been shown to remodel transcription by

regulating heat shock factor 1 (ref. [24]). We checked whether HSF1 integrates mitochondrial signals to activate the UPR^{mt}. Indeed, depletion of *HSF1* abrogated UPR^{mt} induction (Fig. 3e and Extended Data Fig. 7a–c). Strikingly, basal mitochondrial chaperone protein levels were also reduced in *HSF1* KO cells (Extended Data Fig. 7a), suggesting that HSF1 serves as a constitutive key regulator of mitochondrial chaperone transcription. This hypothesis is in line with a previous finding, which showed that HSF1 mediates mitochondrial chaperone expression during mitochondrial stress[25]. In addition to the dependency on HSF1 expression, we found activation of HSF1 during GTPP treatment, monitored by its translocation from the cytosol to the nucleus (Fig. 3f,g). Notably, transcription of non-UPR^{mt}-related mitochondrial proteins was not controlled by HSF1, indicating a separate regulation of the UPR^{mt} and general mitochondrial biogenesis (Extended Data Fig. 7d,e).

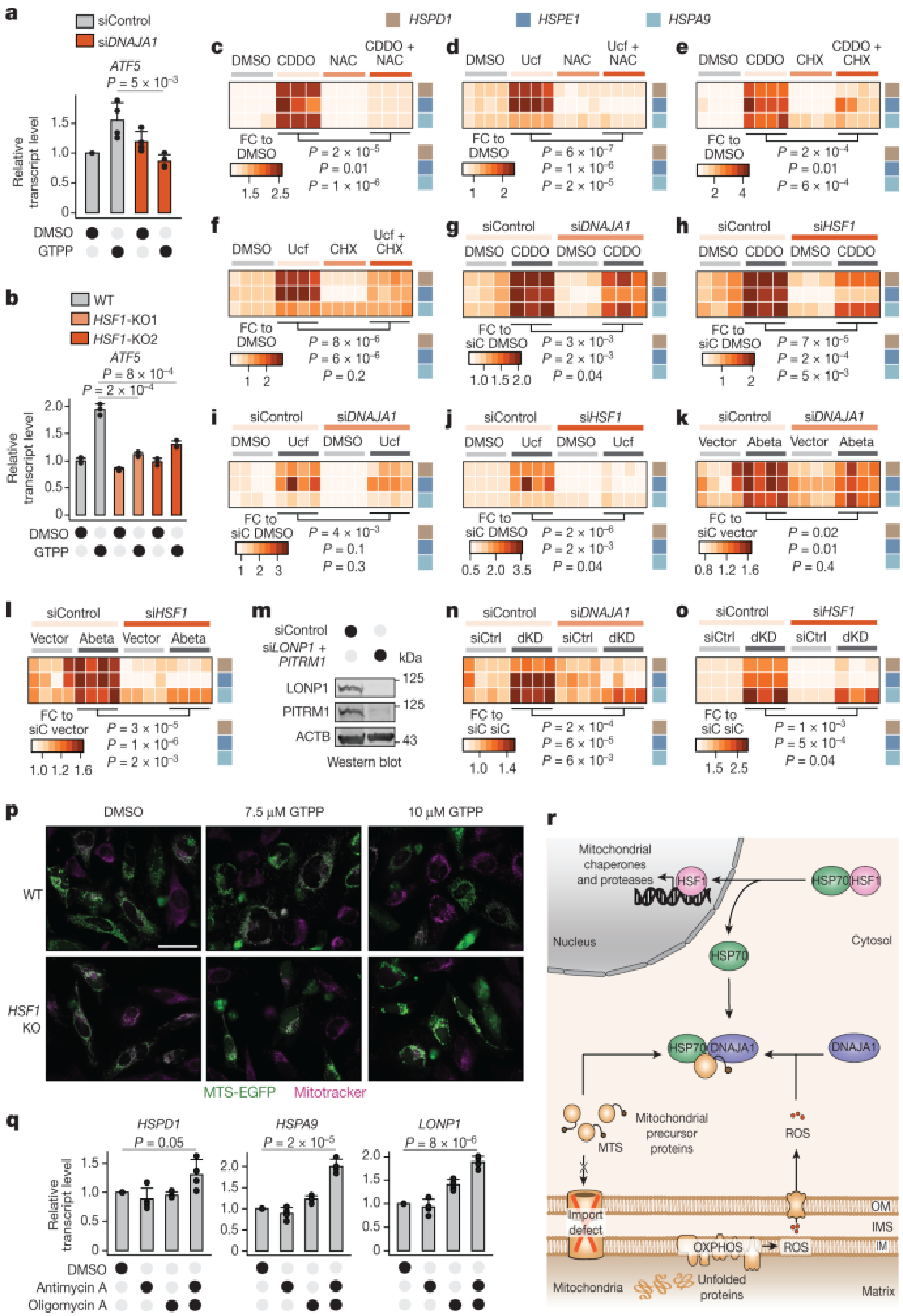
We next tested whether HSF1 activation during MMS requires mtROS and c-mtProt accumulation. Inhibition of ROS signalling by antioxidants (Figs. 1d and 3f,g and Extended Data Fig. 7f), reduction of c-mtProt via cotreatment with CHX (Extended Data Fig. 7g,h) or knockdown of *NRF1* (Extended Data Fig. 7i,j) prevented HSF1 translocation and the UPR^{mt}. These findings indicate that HSF1 activation might take part in the signalling cascade to activate the UPR^{mt} via DNAJA1. To test this hypothesis, we checked whether DNAJA1 was required for the HSF1 activation observed upon MMS. Indeed, depletion of *DNAJA1* significantly inhibited HSF1 translocation to the nucleus (Fig. 3h,i).

Under basal conditions, HSP70 binds to HSF1 and represses its transcriptional activity[26]. Immunoprecipitation experiments of HSF1 showed that the HSF1–HSP70 interaction decreased upon GTPP treatment (Extended Data Fig. 7k,l), suggesting that the recruitment of HSP70 to c-mtProt via DNAJA1 titrates HSP70 away from HSF1. This then leads to HSF1 activation and its subsequent translocation to the nucleus to activate the UPR^{mt}. Overall, our findings define HSF1 as the transcription factor downstream of DNAJA1 that responds to mtROS and c-mtProt accumulation to induce the UPR^{mt}.

The DNAJA1–HSF1 axis activates the UPR^{mt}

Here, we found a specific pathway that signals the UPR^{mt} driven by a cytosolic surveillance mechanism that conveys information on both mtROS and c-mtProt accumulation to the nucleus via DNAJA1 and HSF1. While we did not identify the ISR–ATF4 axis to be essential for UPR^{mt} signalling (Extended Data Fig. 1e–h), previous work has shown a potential involvement of the ISR target gene *ATF5* (ref. [6]). However, depletion of *ATF5* did not inhibit UPR^{mt} activation (Extended Data Fig. 8a–e). Intriguingly, *ATF5* transcription increased after UPR^{mt} activation (Extended Data Fig. 8f) and depended on both DNAJA1 and HSF1 (Fig. 4a,b). Thus, we suggested that ATF5 may largely act downstream after UPR^{mt} activation by the cytosolic surveillance. Indeed, analyses of available HSF1 chromatin immunoprecipitation and sequencing (CHIP–seq) data showed that HSF1 binds the ATF5 regulatory region (Extended Data Fig. 8g), and both HSF1 and ATF5 are able to bind directly to the promoters of mitochondrial chaperones (Extended Data Fig. 8h). Thus, while ATF5 was not required for UPR^{mt} activation, it might fulfil essential functions later.

Fig. 4: Different stressors activate the UPR^{mt} via the cytosolic surveillance mechanism.



a
siControl
siDNAJA1
ATF5
P = 5 × 10⁻³
Relative transcript level
DMSO
GTPP
b
WT
HSF1-KO1
HSF1-KO2
ATF5
P = 8 × 10⁻⁴
P = 2 × 10⁻⁴
HSPD1
HSPE1
HSPA9
c
DMSO CDDO NAC CDDO + NAC
FC to DMSO
P = 2 × 10⁻⁵
P = 0.01
P = 1 × 10⁻⁶
d
DMSO Ucf NAC Ucf + NAC
FC to DMSO
P = 6 × 10⁻⁷
P = 1 × 10⁻⁶
P = 2 × 10⁻⁵
e
DMSO CDDO CHX CDDO + CHX
FC to DMSO
P = 2 × 10⁻⁴
P = 0.01
P = 6 × 10⁻⁴
f
DMSO Ucf CHX Ucf + CHX
FC to DMSO
P = 8 × 10⁻⁶
P = 6 × 10⁻⁶
P = 0.2
g
siControl siDNAJA1
DMSO CDDO DMSO CDDO
FC to siC DMSO
P = 3 × 10⁻³
P = 2 × 10⁻³
P = 0.04
h
siControl siHSF1
DMSO CDDO DMSO CDDO
FC to siC DMSO
P = 7 × 10⁻⁵
P = 2 × 10⁻⁴
P = 5 × 10⁻³
i
siControl siDNAJA1
DMSO Ucf DMSO Ucf
FC to siC DMSO
P = 4 × 10⁻³
P = 0.1
P = 0.3
j
siControl siHSF1
DMSO Ucf DMSO Ucf
FC to siC DMSO
P = 2 × 10⁻⁶
P = 2 × 10⁻³
P = 0.04
k
siControl siDNAJA1
Vector Abeta Vector Abeta
FC to siC vector
P = 0.02
P = 0.01
P = 0.4
l
siControl siHSF1
Vector Abeta Vector Abeta
FC to siC vector
P = 3 × 10⁻⁵
P = 1 × 10⁻⁶
P = 2 × 10⁻³
m
siControl
siLONP1 + PITRM1
kDa
LONP1
PITRM1
ACTB
Western blot
n
siControl siDNAJA1
siCtrl dKD siCtrl dKD
FC to siC siC
P = 2 × 10⁻⁴
P = 6 × 10⁻⁵
P = 6 × 10⁻³
o
siControl siHSF1
siCtrl dKD siCtrl dKD
FC to siC siC
P = 1 × 10⁻³
P = 5 × 10⁻⁴
P = 0.04
p
DMSO
7.5 µM GTPP
10 µM GTPP
WT
HSF1 KO
MTS-EGFP
Mitotracker
q
HSPD1
P = 0.05
HSPA9
P = 2 × 10⁻⁵
LONP1
P = 8 × 10⁻⁶
Relative transcript level
DMSO
Antimycin A
Oligomycin A
r
Mitochondrial chaperones and proteases
HSF1
HSP70
Nucleus
Cytosol
DNAJA1
Mitochondrial precursor proteins
MTS
ROS
Import defect
OXPHOS
OM
IMS
IM
Mitochondria
Unfolded proteins
Matrix

a,**b**, Bar plots showing the mean of relative transcript levels of *ATF5* upon *DNAJA1* knockdown ($n = 4$ biological replicates) (**a**) and *HSF1* KO ($n = 3$ biological replicates) (**b**) measured by qPCR. **c**,**d**, Heat maps of relative transcript levels of UPR^{mt} genes in CDDO- (**c**) or Ucf-101-treated (**d**) cells upon cotreatments with NAC ($n = 3$ and 4 biological replicates for CDDO and Ucf-101, respectively). **e**,**f**, Heat maps of relative transcript levels of UPR^{mt} genes in CDDO- (**e**) or Ucf-101-treated (**f**) cells upon cotreatments with CHX ($n = 4$ biological replicates). **g**–**j**, Heat maps of relative transcript levels of UPR^{mt} genes in *DNAJA1* knockdown cells treated with CDDO (**g**) or Ucf-101 (**i**), or *HSF1* knockdown cells treated with CDDO (**h**) or Ucf-101 (**j**) ($n = 3$ and 4 biological replicates for CDDO and Ucf-101, respectively) (**g–j**) measured by qPCR. siC, siControl. **k**,**l**, Heat maps of relative transcript levels of UPR^{mt} genes upon knockdown of *DNAJA1* (**k**) and *HSF1* (**l**) in cells overexpressing MTS-Abeta ($n = 4$ biological replicates) measured by qPCR. **m**, Western blot image of double knockdown of *LONP1* and *PITRM1* ($n = 2$ biological replicates). **n**,**o**, Heat maps of relative transcript levels of UPR^{mt} genes upon knockdown of *DNAJA1* (**n**) and *HSF1* (**o**) in cells within an *LONP1* and *PITRM1* double knockdown (dKD) background ($n = 4$ biological replicates) measured by qPCR. Transcript levels are represented as relative FCs. **p**, Representative microscopy images of MTS-EGFP (green) localization in comparison with mitochondria (magenta) in wild-type (WT) and *HSF1* KO cells ($n = 5$ biological replicates). Scale bar, 50 µm. **q**, Bar plots showing the mean of relative transcript levels of UPR^{mt} genes measured with qPCR ($n = 4$ biological replicates). **r**, Working model of the cytosolic UPR^{mt} surveillance mechanism and transcriptional UPR^{mt} activation. OM, outer membrane; IM, inner membrane. All *P* values are calculated with a two-tailed unpaired Student's *t* test and indicated. All error bars represent mean ± s.d. Gel source data are in Supplementary Fig. 1.

Source data

We next assessed whether the identified UPR^{mt} cytosolic surveillance is the general mechanism for UPR^{mt} activation upon MMS. Activation of the UPR^{mt} upon inhibition of two different mitochondrial proteases—LON protease (LONP) with 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO) and HTRA2 with dihydro-5-[[5-(2-nitrophenyl)-2-furanyl]methylene]-1,3-diphenyl-2-thioxo-4,6(1H,5H)-pyrimidinedione

(Ucf-101)—was also dependent on ROS and accumulation of c-mtProt (Fig. 4c–f), and it was mediated by DNAJA1 and HSF1 (Fig. 4g–j). Thus, signalling across the ROS + c-mtProt–DNAJA1–HSF1 axis is a common pathway used for UPR^{mt} activation in human cells. Next, we induced the UPR^{mt} genetically by overexpression of the aggregation-prone protein Abeta in mitochondria or double knockdown of the mitochondrial proteases *LONP1* and *PITRM1* and found that it was also dependent on DNAJA1 and HSF1 (Fig. 4k–o and Extended Data Fig. 9a,b). These observations underline the relevance of the cytosolic surveillance mechanism for the maintenance of mitochondrial proteostasis in a physiological context. Indeed, knocking out *HSF1* increased mitochondrial vulnerability upon MMS, leading to import defects and reduced overall cell survivability (Fig. 4p and Extended Data Fig. 9c–f).

Finally, we evaluated whether ROS + c-mtProt can directly activate the UPR^{mt} signalling cascade without upstream induction of MMS. Employing oxidative phosphorylation (OXPHOS) inhibitors antimycin A or oligomycin A individually was not sufficient to simultaneously induce mtROS and c-mtProt and to activate the UPR^{mt}(refs. [27,28]) (Extended Data Fig. 10a–h). However, the combination of antimycin A and oligomycin A increased ROS and c-mtProt, and it induced the UPR^{mt} without mitochondrial protein aggregate formation (Fig. 4q and Extended Data Fig. 10a–i). In addition, inspecting different mitochondrial stressors confirmed that individual induction of ROS or c-mtProt accumulation alone did not induce the UPR^{mt} (Extended Data Fig. 10j–l). These findings support that mitochondrial stress can activate UPR^{mt} signalling only when both ROS and c-mtProt accumulation are induced.

Discussion

Our data uncover the signalling molecules used by mitochondria to initiate the UPR^{mt}, a two-pronged signalling cascade composed of mtROS and c-mtProt (Fig. 4r). mtROS and c-mtProt signals converge at the DNAJA1-mediated activation of HSF1, forming a surveillance mechanism in the cytosol to initiate the transcriptional programme of mitochondrial chaperones and proteases. These findings reveal an unexpected connection

between the mitochondrial and cytosolic proteostasis networks during UPR^{mt} activation. Intriguingly, while a canonical UPR^{mt} has not been described in yeast, c-mtProt has been shown to cause remodelling of cytosolic proteostasis[19,20,21,24,29]. In *Caenorhabditis elegans*, perturbation of mitochondrial protein homoeostasis had also been shown to activate cytosolic proteostasis coordinated by lipid biosynthesis[30]. Human cells appear to have evolved these principles to a complex cytosolic surveillance system for UPR^{mt} activation, which links mitochondrial proteostasis to a broader network of cellular homoeostasis. Ultimately, this pathway might provide explanations for diseases in which the breakdown of cytosolic proteostasis is linked to mitochondrial dysfunction, including ageing.

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Cell culture and treatments

HeLa ovarian carcinoma cells from the American Type Culture Collection were used for all experiments unless stated otherwise. They were confirmed to be mycoplasma negative and grown in RPMI medium (Thermo Fisher Scientific) and 10% fetal bovine serum. Knockdown experiments were performed with Lipofectamine RNAiMAX according to the manufacturer’s instructions. The small interfering RNAs (siRNAs) used were from OriGene oligo duplex *ATF5* (SR307793) and custom made for *LONP1* (sense 5′-GGACGUCCUGGAAGAGACCAAUAUU-3′, anti-sense 5′-AAUAUUGGUCUCUUCCAGGACGUCC). MISSION esiRNA (Sigma) were *ATF5* (EHU039491), *DNAJA1* (EHU114481), *HSF1* (EHU107721), *DNAJA2* (EHU005311), *DNAJB1* (EHU109151), *NRF1* (EHU069871) and *PITRM1* (EHU011041). Gene KOs were conducted by CRISPR–Cas9-mediated genome editing. The single guide RNAs (sgRNAs) were cloned into eSpCas9 (1.1; Addgene, catalogue no. 71814). The sgRNA sequences

used were 5′-GCAACAGAAAGTCGTCAACA-3′ (*HSF1*), 5′-TCTCTTAGATGATTACCTGG-3′ (*ATF4*), TCAGCCAAGCCAGAGAAGCA-3′, 5′-ATTTCCAGGAGGTGAAACAT-3′ (*DDIT3*), 5′-TGGCTCCCTATGAGGTCCTT-3′ (*ATF5_1*) and 5′-AGACTATGGGAAACTCCCCC-3′ (*ATF5_2*). Together with sgRNA-containing plasmid, cells were cotransfected with puromycin-resistant plasmids and selected for 24 h with 1 μg ml^{-1} puromycin (Invivogen). After the selection, single cells were seeded into 96-well plates and incubated for 2 weeks. Resulting colonies were expanded, and gene KO was confirmed by Sanger sequencing and western blot.

Transient overexpression of MTS-Abeta-GFP was carried out with Lipofectamine 2000 according to the manufacturer's instruction. Cells were harvested after 24 h.

Acute induction of the UPR^{mt} was performed with 10 μM GTPP (Shanghai Chempartner), 5 μM CDDO (Cayman Chemical) and 40 μM Ucf-101 (Cayman Chemical) for 6 h unless stated otherwise (for early response, a 3 h incubation was used). For the cell viability assay, a toxic concentration of 15 μM GTPP for 16 h was applied (Extended Data Fig. 9d–f). mtROS induction was done by treating cells with 10 μM antimycin A (Sigma) or 2 μM rotenone (Sigma) for 6 h. To scavenge ROS, cells were pretreated with 10 mM NAC (Sigma) or 10 mM GSH (Cayman Chemical) for 1 h or 100 μM MnTBAP (Sigma) overnight that was continued as a cotreatment. For Hyper7 references, 20 μM antimycin A and 1 mM H_2O_2 (Carl Roth) were used; 4,4′diisothiocyanatostilbene-2,2′-disulfonate (75 μM, Sigma) was used to inhibit VDAC1 for 6 h. General translation was blocked by treatment with 35 μM CHX for 30 min and continued as cotreatment for 6 h. Mitochondrial import inhibition was performed with 5 μM oligomycin A (Sigma) for 6 h. Different mitochondrial stressors were applied by 6 h of treatment with 10 μM carbonyl cyanide m-chlorophenyl hydrazone (CCCP, Abcam), 100 μM deferiprone (DFP, Sigma) or 10 μM Menadione (Sigma). The hypoxic condition was generated by incubating cells in the BD GasPak EZ Pouch system (BD Diagnostics) for 6 h. Staurosporine (Cayman Chemical) was used to induce apoptosis as a control treatment at 1 μM for 3 h or 200 nM overnight for the *HSF1* KO cell viability assay (Extended Data Fig. 9d–f).

Cloning

For the generation of the construct MTS-Abeta-GFP, pcDNA5/FRT/TO (Thermo) was used as a backbone. The following inserts were amplified by Q5 High-Fidelity DNA Polymerase (NEB) and cloned into the backbone via NEBuilder HiFi DNA Assembly Master Mix (NEB): MTS (2× COX8 presequence in tandem) amplified from pCMV CEPIA2mt (Addgene, catalogue no. 58218), Abeta (Aβ1-42) amplified from HeLa wild-type complementary DNA (cDNA) with primers (5′-TCC ATG CGG GGT TCT GAT GCA GAA TTC CGA CAT GAC TCA GGA TAT G-3′ and 5′-CTC GCC CTT GCT CAC GGA TCC CGC TAT GAC AAC ACC GCC CAC C-3′, containing a GS linker) and enhanced green fluorescent protein (EGFP) amplified from Su9-EGFP (Addgene, catalogue no. 23214).

RNA sequencing

Total RNAs were extracted from cells using the NucleoSpin RNA Plus kit (Macherey-Nagel) following the manufacturer's instructions and subsequently digested with Turbo DNase (Thermo Fisher Scientific). Library preparation for bulk sequencing of poly(A)-RNA was done as described previously[31]. Briefly, barcoded cDNA of each sample was generated with a Maxima RT polymerase (Thermo Fisher Scientific) using an oligo-dT primer containing barcodes, unique molecular identifiers (UMIs) and an adaptor. Ends of the cDNAs were extended by a template switch oligo, and full-length cDNA was amplified with primers binding to the template switch oligo site and the adaptor. The NEB UltraII FS kit was used to fragment cDNA. After end repair and A tailing, a TruSeq adaptor was ligated, and 3′-end fragments were finally amplified using primers with Illumina P5 and P7 overhangs. In comparison with Parekh et al.[31], the P5 and P7 sites were exchanged to allow sequencing of the cDNA in read1 and barcodes and UMIs in read2 to achieve a better cluster recognition. The library was sequenced on a NextSeq 500 (Illumina) with 63 cycles for the cDNA in read1 and 16 cycles for the barcodes and UMIs in read2.

RNA sequencing analysis

Gencode gene annotations v.35 and the human reference genome GRCh38 were derived from the Gencode homepage (European Molecular Biology's European Bioinformatics Institute (EMBL-EBI)). Drop-Seq tools (v.1.12)[32] were used for mapping raw sequencing data to the reference genome. The resulting UMI filtered count matrix was imported into R (v.4.0.5), and lowly expressed genes were subsequently filtered out. Data were then variance stabilized via the rlog function as implemented in DESeq2 (v.1.18.1)[33]. For accurate dispersion estimation, the experimental design (treatment at a given time point) was provided to the function. rlog normalized data were used to perform clustering analysis (fuzzy C means) with R package mFuzz (v.2.50.0)[34]. Transcripts with rlog normalized values of less than three were excluded from the analysis. The number of clusters was set to three. Transcripts were assigned to increased and decreased cluster groups based on cluster membership greater than or equal to 0.8 for each cluster. Gene Ontology (GO) enrichment analysis was performed on each cluster group by using Database for Annotation, Visualization and Integrated Discovery (DAVID). GO enrichments were visualized with the EnrichmentMap (v.3.3.2) plug-in in Cytoscape (v.3.7.1).

Quantitative polymerase chain reaction analysis

Total RNAs were extracted from cells using the NucleoSpin RNA Plus kit (Macherey-Nagel) following the manufacturer's instructions. cDNA synthesis was performed with the High-capacity cDNA reverse transcription kit (Applied Biosystems). Quantitative polymerase chain reaction (qPCR) analysis was performed with primaQuant SYBRGreen master mix without ROX (Steinbrenner Laborsysteme) according to the manufacturer's instructions. KiCqStart primers SYBR green from Sigma (Supplementary Table 4) were used to perform qPCR measurement with LightCycler 480 SW (v.1.5) on the LightCycler 480 real-time PCR system (Roche) in 384-well format. *ACTB* was used as an internal control. Fold changes of the transcript level were calculated using the comparative CtΔΔCt (cycle threshold) method.

FACS measurement

MitoSOX Red (Thermo Fisher Scientific) was used to measure mtROS production according to the manufacturer's instructions. Cell deaths were measured with a combination of Annexin V conjugated to Alexa Fluor 488 (Thermo Fisher Scientific) and propidium iodide (Thermo Fisher Scientific) according to the manufacturer's instructions. Fluorescence-activated cell sorting (FACS) was performed with FACSDiva (v.6.1.3) on FACSCanto II and FACSymphony A5 flow cytometry systems (BD) for MitoSOX and Annexin V measurements, respectively. Analysis of FACS data was performed with FlowJo v.10 software.

Immunoblotting

Cells were lysed in RIPA buffer containing Complete Mini EDTA-free protease inhibitor (Roche) and GENIUS nuclease (Santa Cruz Biotechnology). Lysates were prepared in 1× Laemmli buffer and boiled for 10 min at 95 °C. Proteins were separated with SDS–PAGE using the Invitrogen Novex system and transferred to nitrocellulose membrane by using Mini Trans-Blot cell (Bio-Rad). Primary antibodies were added to immunoblots for 1 h at room temperature (RT). Antibodies used for the detection were anti-ACTB (SantaCruz, catalogue no. sc69879, 1:4,000), anti-HSPD1 (Abcam, catalogue no. ab46798, 1:2,000), anti-COX5B (Proteintech, catalogue no. 11418-2-AP, 1:1,000), anti-HSF1 (Cell Signaling, catalogue no. 4356, 1:1,000), anti-HSF1 (Abcam, catalogue no. ab2923, 1:10,000), anti-DNAJA1 (Proteintech, catalogue no. 11713-1-AP, 1:2,000), anti-Hsp70 (Proteintech, catalogue no. 10995-1-AP, 1:2,000), anti-NRF1 (Cell Signaling (D9K6P), catalogue no. 46743, 1:1,000), anti-α-tubulin (Cell Signaling (DM1A), catalogue no. 3873, 1:3,000), anti-histone H3 (Active Motif, catalogue no. 39163, 1:5,000), anti-FLAG (Sigma, catalogue no. F1804, 1:5,000), anti-CHOP (Thermo Fisher Scientific, catalogue no. MA1-250, 1:1,000), anti-ATF4 (Cell Signaling, catalogue no. 11815, 1:1,000), anti-PITRM1 (Novus, catalogue no. H00010531-M03, 1:500), anti-LONP1 (Proteintech, catalogue no. 15440-1-AP, 1:2,000), anti-cleaved PARP1 (Cell Signaling, catalogue no. 5625, 1:2,000) and anti-Caspase3 (Cell Signaling, catalogue no. 9661, 1:1,000). Secondary antibodies used were anti-rabbit IgG (H + L) HRP Conjugate (Promega, catalogue no. W4021, 1:10,000), IRDye 800CW goat anti-rabbit IgG (H + L; Li-Cor, catalogue no. 926–32211, 1:15,000) and IRDye 680RD donkey anti-

mouse IgG (H + L; Li-Cor, catalogue no. 926–68072, 1:15,000). Appropriate secondary antibodies were used for imaging with Odyssey DLx (LI-COR) or ChemiDoc MP (Bio-Rad) imaging system. Data were collected with Image Studio (v.5.2) or ImageLab v.6.0.1.

Mitochondrial insoluble fraction analysis

Mitochondrial fractions were prepared as previously described[35]. Briefly, cells were homogenized by passing them through a 27-gauge needle syringe in buffer containing 10 mM HEPES (pH 7.4), 50 mM sucrose, 0.4 M mannitol, 10 mM KCl and 1 mM EGTA. Mitochondrial enrichment was performed with a two-step differential centrifugation at 1,000*g* followed by 13,000*g* for 15 min each at 4 °C. The mitochondria-enriched pellets were resuspended in a buffer containing 20 mM HEPES (pH 7.4), 0.4 M mannitol, 10 mM NaH_2PO_4 and 0.5 M EGTA. An equal volume of lysis buffer containing 2% (vol/vol) NP40 was added and spun down to separate mitochondrial fractions. The resulting supernatants and pellets were kept as the soluble and insoluble fractions, respectively. Proteins were resolved with SDS–PAGE in 1× Laemmli buffer and visualized with InstantBlue Coomassie stain (Expedeon).

Nuclear and cytosolic fractionation

Cells were fractionated with the REAP method[36]. Cell fractions were prepared by resuspending cells in PBS containing 0.1% (vol/vol) NP40, followed by five times resuspension with a p1000 micropipette (Gilson). Cells were fractionated with a ‘pop spin’ for 10 s at 4 °C in an Eppendorf tabletop microfuge. Supernatants were collected as the cytosolic fractions. Pellets were washed once with 0.1% (vol/vol) NP40 and collected as the nuclear fractions. Both the cytoplasmic and nuclear fractions were used to perform immunoblotting. The ratio of nuclear to cytosolic HSF1 was calculated as follows:

HSF1 (N/C) = (Nuclear HSF1/Histone H3)/(Cytoplasmic HSF1/Tubulin).

Immunoprecipitation

Crosslinking was performed by incubating cells in PBS containing 0.8 mg ml^{-1} dithiobis[succinimidyl propionate] (Proteochem) for 30 min at RT[37]. Crosslinking reactions were quenched with PBS containing 200 µM glycine for 15 min at RT. Cells were lysed in cell lysis buffer (50 mM Tris (pH 8.0), 150 mM NaCl, 1% (vol/vol) NP40) containing protease inhibitor and allowed to incubate for 30 min at 4 °C. Lysates containing 2 mg of total proteins were used to perform immunoprecipitation with 10 µl Dynabeads protein A (Thermo Fischer Scientific) containing 1 µg of appropriate antibodies or 10 µl Anti-FLAG M2 magnetic beads (Sigma) for 2 h at 4 °C. Immunoprecipitated proteins were eluted from beads for immunoblotting or digested for interaction proteomics.

Sample preparation for LC–MS/MS

For redox proteomics, cells were lysed in HES buffer (1 mM EDTA, 0.1% (wt/vol) SDS, 50 mM HEPES (pH 8.0)) supplemented with protease inhibitor and 10% (vol/vol) TCA and incubated for 2 h at 4 °C. Each sample was divided into two fractions: (1) oxidized Cys fraction and (2) total Cys fraction. Proteins were precipitated with a TCA and acetone precipitation. For fraction 2, 100 µg of proteins were resuspended in HES buffer supplemented with 5 mM TCEP and incubated for 1 h at 50 °C to reduce all Cys thiols. For fraction 1, 100 µg of the proteins were resuspended in denaturing buffer (6 M urea, 1% (wt/vol) octyl ß-glucopyranoside, 50 mM HEPES (pH 8.0)) supplemented with protease inhibitor and 200 mM iodoacetamide and incubated for 1 h at 37 °C in the dark to block free Cys thiols. Oxidized Cys thiols were reduced as described previously for fraction 2. Proteins were cleaned up by TCA and acetone precipitation. To label the free Cys thiols, proteins were resuspended in denaturing buffer supplemented with iodoTMT#1 (Thermo Fisher Scientific) for fraction 1 or iodoTMT#2 for fraction 2 and incubated for 1 h at 37 °C in the dark. Labelling reactions were quenched with 20 mM DTT. Labelled proteins were pooled together and cleaned up with TCA and acetone precipitation. Proteins were digested with 1:50 (wt/wt) LysC (Wako Chemicals) and 1:100 (wt/wt) Trypsin (Promega) in 10 mM EPPS (pH 8.2) containing 1 M urea overnight at 37 °C. Peptides were purified with (50-mg) SepPak columns (Waters) and then dried. IodoTMT-labelled peptides were enriched with anti-TMT antibody resin (Thermo Fisher Scientific) according to the

manufacturer’s instructions. Enriched pools of labelled peptides were subjected to high-pH reverse-phase fractionation with the High pH RP Fractionation kit (Thermo Fisher Scientific) following the manufacturer’s instructions. Fractionated peptides were concatenated into four separate fractions.

To perform interaction proteomics, after immunoprecipitation steps 25 µl of SDC (2% SDC (wt/vol), 1 mM TCEP, 4 mM chloroacetamide, 50 mM Tris (pH 8.5)) buffer was added to the beads. The mixtures were heated up to 95 °C, and the supernatants were collected. For digestion, 25 µl of 50 mM Tris (pH 8.5) containing 1:50 (wt/wt) LysC (Wako Chemicals) and 1:100 (wt/wt) trypsin (Promega) was added and allowed to incubate overnight at 37 °C. Digestion was stopped by adding 150 µl of isopropanol containing 1% (vol/vol) TFA. Peptide purification was performed with the SDB-RPS disc (Sigma) and then dried.

LC–MS/MS

Peptides were resuspended in a 2% (vol/vol) acetonitrile/1% (vol/vol) formic acid solution and separated on an Easy nLC 1200 (Thermo Fisher Scientific) and a 35-cm-long, 75-µm-inner-diameter fused-silica column, which had been packed in house with 1.9-µm C18 particles (ReproSil-Pur, Dr. Maisch) and kept at 50 °C using an integrated column oven (Sonation). For redox proteome, peptides were eluted by a nonlinear gradient from 4 to 36% (vol/vol) acetonitrile over 90 min and directly sprayed into a QExactive HF mass spectrometer equipped with a nanoFlex ion source (Thermo Fisher Scientific) at a spray voltage of 2.3 kV. Full-scan MS spectra (350–1,400 *m*/*z*) were acquired at a resolution of 120,000 at *m*/*z* 200, a maximum injection time of 25 ms and an automatic gain control (AGC) target value of 3×10^6. Up to 20 of the most intense peptides per full scan were isolated using a 1-Th window and fragmented using higher-energy collisional dissociation (normalized collision energy of 35). MS/MS spectra were acquired with a resolution of 45,000 at *m*/*z* 200, a maximum injection time of 86 ms and an AGC target value of 1×10^5. Ions with charge states of one, five to eight and more than eight as well as ions with unassigned charge states were not considered for fragmentation. Dynamic exclusion was set to 20 s to minimize repeated sequencing of already acquired precursors.

For interaction proteomics, peptides were eluted by a nonlinear gradient from 3.2 to 32% acetonitrile over 60 min followed by a stepwise increase to 95% B in 6 min, which was kept for another 9 min and sprayed into an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific) at a spray voltage of 2.3 kV. Full-scan MS spectra (350–1,500 *m*/*z*) were acquired at a resolution of 60,000 at *m*/*z* 200, a maximum injection time of 50 ms and an AGC target value of 4×10^5. The most intense precursors with a charge state between two and six per full scan were selected for fragmentation ('Top Speed' with a cycle time of 1.5 s) and fragmented using higher-energy collisional dissociation (normalized collision energy of 30). MS/MS spectra were acquired with a resolution of 15,000 at *m*/*z* 200, a maximum injection time of 22 ms and an AGC target value of 1×10^5. Ions with charge states of one and more than six as well as ions with unassigned charge states were not considered for fragmentation. Dynamic exclusion was set to 45 s to minimize repeated sequencing of already acquired precursors.

LC–MS/MS data analysis

For analysis of redox proteomics data, raw files were analysed using Proteome Discoverer 2.4 software (Thermo Fisher Scientific). Spectra were selected using default settings and database searches performed using the SequestHT node in Proteome Discoverer. Database searches were performed against a trypsin-digested *Homo sapiens* SwissProt database and FASTA files of common contaminants ('contaminants.fasta' provided with MaxQuant) for quality control. Dynamic modifications were set as methionine oxidation (C, +15.995 Da), iodoTMT6plex (C, +329.227 Da) and carbamidomethyl (C, +57.021 Da) at cysteine residues. One search node was set up to search with Met loss + acetyl (M, −89.030 Da) as dynamic modifications at the N terminus. Searches were performed using Sequest HT. After each search, posterior error probabilities were calculated, and peptide spectrum matches were filtered using Percolator with default settings. Consensus workflow for reporter ion quantification was performed with default settings, except that the minimal signal-to-noise ratio was set to 10. Results were then exported to Excel files for further processing. Non-normalized abundances were used for quantification. The percentage of cysteine oxidation for each peptide was calculated as follows:

Percentage of oxidized Cys = (abundance of fraction 1/abundance of fraction 2) × 100%.

For peptides with several different Cys modifications, fold changes of the percentage of oxidized Cys from each different combination were considered.

For DNAJA1 interaction proteomics, MS raw data processing was performed with MaxQuant (v.1.6.17.0) and its in-build label-free quantification algorithm MaxLFQ applying default parameters[38]. Acquired spectra were searched against the human reference proteome (Taxonomy identification 9606) downloaded from UniProt (12-03-2020; 'One sequence per gene', 20,531 sequences) and a collection of common contaminants (244 entries) using the Andromeda search engine integrated in MaxQuant[39]. Identifications were filtered to obtain false discovery rates below 1% for both peptide spectrum matches (minimum length of seven amino acids) and proteins using a target-decoy strategy[40]. Results were then exported to Excel files for further processing. Abundance of interactors was normalized to the abundance of DNAJA1 from each sample. Fold changes were calculated from normalized data. GO enrichment analysis of DNAJA1 interactome was performed by using DAVID. GO enrichments were visualized with the EnrichmentMap (v.3.3.2) plug-in in Cytoscape (v.3.7.1). Subcellular locations of increased interactors upon GTPP treatment were manually curated from UniProt.

Microscopy analysis

For HyPer7 measurements, cells were transfected with different constructs of HyPer7 (ref. [13]) (Addgene, catalogue nos. 136466, 136469 and 136470) with Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Measurements were performed 24 h after transfection in a 96-well plate format at 37 °C. Time-series live cell imaging was done with a CQ1 confocal imaging cytometer (Yokogawa). HyPer7 was excited sequentially with 405- and 488-nm laser beams. Emission was collected using a 525/50-bandpass emission filter. After five images were acquired, 10 μM GTPP was added to each group of cells expressing different constructs. Image analysis was performed using ImageJ (v.1.53).

Fluorescence was calculated for regions of interests inside the imaged cell. The ratiometric signal of HyPer7 was calculated by dividing the intensity of the emission signals excited by 488/405 nm.

Monitoring of DNAJA1 localization was performed using SP8 Confocal (Leica). Cells were incubated in media containing 150 nM MitoTracker deep red FM (Thermo Fisher Scientific) for 30 min at 37 °C in the dark. After subsequent washes, the cells were fixed with 4% (vol/vol) formaldehyde in PBS and permeabilized with 0.01% (vol/vol) TritonX100. Cells were blocked with a PBS buffer containing 1% (wt/vol) bovine serum albumin (BSA), 300 mM glycine and 0.1% (vol/vol) Tween20 for 30 min at RT. Cells were incubated overnight at 4 °C with 1:100 dilutions of anti-DNAJA1 antibody (11713-1-AP, Proteintech) in PBS containing 1% (wt/vol) BSA and 0.1% (vol/vol) Tween20. After 3× washes, 1:1,000 dilution of Alexa Fluor 488 anti-rabbit IgG in 1% BSA PBS was used to incubate the cells for 1 h at RT as a secondary antibody. A drop of ProLong diamond antifade mountant containing DAPI (Thermo Fisher Scientific) was used to mount the cells. Data were collected with Leica Application Suite X. Image analysis was performed using ImageJ (v.1.53). Pearson's and Mander's colocalization coefficients were calculated with the JACoP plug-in[41]. Calculation from the independent images was reported.

For monitoring mitochondrial import, MTS-EGFP (Addgene, catalogue no. 23214) was transiently transfected together with 10 μM GTPP treatment. After 6 h of incubation, cells were stained with 50 nM Mitotracker Deep Red FM (Thermo Fisher Scientific) for 15 min and transferred to RPMI 10% fetal calf serum for live cell imaging. CQ1 (Yokogawa) with 40× magnification was used with the following laser settings: 488-nm excitation and 525/50-nm emission for EGFP and 640-nm excitation and 685/40-nm emission for Mitotracker Deep Red FM. Three wells with a total of 100 cells per condition were manually characterized into one of five categories.

For Halo-tagged reporter assay, T-Rex-HeLa cells stably expressing Halo-tagged ATP5A1 and GREPL1 were used. Blocking of previously synthesized Halo-tagged proteins was done by incubating cells in media containing 5 μM empty HaloTag ligand (Promega) overnight. Treatments were started on the following day. Newly synthesized Halo-tagged proteins were labelled with 5 μM HaloTag TMR ligand (Promega) in the last hour of

the treatments. Mitochondria were stained with 50 nM Mitotracker Deep Red FM (Thermo Fisher Scientific). CQ1 (Yokogawa) with 40× magnification was used with following laser settings: 488-nm excitation and 525/50-nm emission for TMR and 561-nm excitation and 617/73-nm emission for Mitotracker Deep Red FM. Several images were collected from three independent replicates, and image analysis was performed using ImageJ (v.1.53). Pearson's and Mander's colocalization coefficients were calculated with the JACoP plug-in[41]. Calculation from three independent replicates was reported.

Cell viability assay

Cell viability was measured with Cell Counting Kit-8 (Dojindo) according to the manufacturer's instructions. Cell viability was evaluated 16 h (overnight) after treatment with the different chemicals used in the experiment.

Statistics and plots

A general statistical analysis was performed with a two-tailed Student's *t* test (considered significant for $P \leq 0.05$), unless it was stated otherwise. All plots were created using the R packages ggplot2 (v.3.3.3), gplots (v.3.1.1) and RColorBrewer (v.1.1-2). Visualization of the final figures was done with Adobe Illustrator CS5.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The transcriptomics data have been deposited to the European Nucleotide Archive at EMBL-EBI under accession number PRJEB61069. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[42] via the PRIDE partner repository[43] with the dataset identifier PXD031948 for the DNAJA1 interaction proteomics and PXD032011 for

the redox proteomics data. HSF1 and ATF5 CHIP–seq data were obtained from the ENCODE[44,45] database (https://www.encodeproject.org/). The HSF1 CHIP–seq dataset accession is ENCSR000EET, and the file accession is ENCFF797ENQ. The ATF5 CHIP–seq dataset accession is ENCSR887TWV, and the file accession is ENCFF638RRU. Datasets representing the key findings of this paper are within the main and supplementary figures and tables of this article. Unprocessed images are available on request from the corresponding author. Source data are provided with this paper.

Code availability

No custom code was used in this study.

References

1. Münch, C. & Harper, J. W. Mitochondrial unfolded protein response controls matrix pre-RNA processing and translation. *Nature* **534**, 710–713 (2016).

2. Zhao, Q. et al. A mitochondrial specific stress response in mammalian cells. *EMBO J.* **21**, 4411–4419 (2002).

3. Münch, C. The different axes of the mammalian mitochondrial unfolded protein response. *BMC Biol.* **16**, 81 (2018).

4. Anderson, N. S. & Haynes, C. M. Folding the mitochondrial UPR into the integrated stress response. *Trends Cell Biol.* **30**, 428–439 (2020).

5. Nargund, A. M., Pellegrino, M. W., Fiorese, C. J., Baker, B. M. & Haynes, C. M. Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation. *Science* **337**, 587–590 (2012).

6. Fiorese, C. J. et al. The transcription factor ATF5 mediates a mammalian mitochondrial UPR. *Curr. Biol.* **26**, 2037–2043 (2016).

7. Quirós, P. M. et al. Multi-omics analysis identifies ATF4 as a key regulator of the mitochondrial stress response in mammals. *J. Cell Biol.* **216**, 2027–2045 (2017).

8. Fessler, E. et al. A pathway coordinated by DELE1 relays mitochondrial stress to the cytosol. *Nature* **579**, 433–437 (2020).

9. Kaspar, S. et al. Adaptation to mitochondrial stress requires CHOP-directed tuning of ISR. *Sci. Adv.* **7**, eabf0971 (2022).

10. Forsström, S. et al. Fibroblast growth factor 21 drives dynamics of local and systemic stress responses in mitochondrial myopathy with mtDNA deletions. *Cell Metab.* **30**, 1040–1054.e7 (2019).

11. Horibe, T. & Hoogenraad, N. J. The Chop gene contains an element for the positive regulation of the mitochondrial unfolded protein response. *PLoS ONE* **2**, e835 (2007).

12. Holmström, K. M. & Finkel, T. Cellular mechanisms and physiological consequences of redox-dependent signalling. *Nat. Rev. Mol. Cell Biol.* **15**, 411–421 (2014).

13. Pak, V. V. et al. Ultrasensitive genetically encoded indicator for hydrogen peroxide identifies roles for the oxidant in cell migration and mitochondrial function. *Cell Metab.* **31**, 642–653.e6 (2020).

14. Sies, H. & Jones, D. P. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nat. Rev. Mol. Cell Biol.* **21**, 363–383 (2020).

15. Faust, O. et al. HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature* **587**, 489–494 (2020).

16. Choi, H.-I. et al. Redox-regulated cochaperone activity of the human DnaJ homolog Hdj2. *Free Radic. Biol. Med.* **40**, 651–659 (2006).

17. Kim, J.-S. et al. DksA-DnaJ redox interactions provide a signal for the activation of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA* **115**, E11780–E11789 (2018).

18. Song, J., Herrmann, J. M. & Becker, T. Quality control of the mitochondrial proteome. *Nat. Rev. Mol. Cell Biol.* **22**, 54–70 (2021).

19. Weidberg, H. & Amon, A. MitoCPR—a surveillance pathway that protects mitochondria in response to protein import stress. *Science* **360**, eaan4146 (2018).

20. Wrobel, L. et al. Mistargeted mitochondrial proteins activate a proteostatic response in the cytosol. *Nature* **524**, 485–488 (2015).

21. Wang, X. & Chen, X. J. A cytosolic network suppressing mitochondria-mediated proteostatic stress and cell death. *Nature* **524**, 481–484 (2015).

22. Schäfer, J. A., Bozkurt, S., Michaelis, J. B., Klann, K. & Münch, C. Global mitochondrial protein import proteomics reveal distinct regulation by translation and translocation machinery. *Mol. Cell* **82**, 435–446.e7 (2022).

23. Michaelis, J. B., Bozkurt, S., Schäfer, J. A. & Münch, C. Monitoring mitochondrial protein import using mitochondrial targeting sequence (MTS)-eGFP. *Bio Protoc.* **12**, e4578 (2022).

24. Boos, F. et al. Mitochondrial protein-induced stress triggers a global adaptive transcriptional programme. *Nat. Cell Biol.* **21**, 442–451 (2019).

25. Katiyar, A. et al. HSF1 is required for induction of mitochondrial chaperones during the mitochondrial unfolded protein response. *FEBS Open Bio* **10**, 1135–1148 (2020).

26. Masser, A. E. et al. Cytoplasmic protein misfolding titrates Hsp70 to activate nuclear Hsf1. *eLife* **8**, e47791 (2019).

27. Wachter, C., Schatz, G. & Glick, B. S. Protein import into mitochondria: the requirement for external ATP is precursor-specific whereas intramitochondrial ATP is universally needed for translocation into the matrix. *Mol. Biol. Cell* **5**, 465–474 (1994).

28. Michaelis, J. B. et al. Protein import motor complex reacts to mitochondrial misfolding by reducing protein import and activating mitophagy. *Nat. Commun.* **13**, 5164 (2022).

29. Nowicka, U. et al. Cytosolic aggregation of mitochondrial proteins disrupts cellular homeostasis by stimulating the aggregation of other proteins. *eLife* **10**, e65484 (2021).

30. Kim, H.-E. et al. Lipid biosynthesis coordinates a mitochondrial-to-cytosolic stress response. *Cell* **166**, 1539–1552.e16 (2016).

31. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. The impact of amplification on differential expression analyses by RNA-seq. *Sci. Rep.* **6**, 25533 (2016).

32. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

33. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

34. Futschik, M. E. & Carlisle, B. Noise-robust soft clustering of gene expression time-course data. *J. Bioinform. Comput. Biol.* **3**, 965–988 (2005).

35. Burbulla, L. F. et al. Mitochondrial proteolytic stress induced by loss of mortalin function is rescued by Parkin and PINK1. *Cell Death Dis.* **5**, e1180–e1180 (2014).

36. Suzuki, K., Bose, P., Leong-Quong, R. Y. Y., Fujita, D. J. & Riabowol, K. REAP: a two minute cell fractionation method. *BMC Res. Notes* **3**, 294 (2010).

37. Wang, H., He, M., Willard, B. & Wu, Q. Cross-linking, immunoprecipitation and proteomic analysis to identify interacting proteins in cultured cells. *Bio Protoc.* **9**, e3258 (2019).

38. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

39. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

40. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214 (2007).

41. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).

42. Vizcaíno, J. A. et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–226 (2014).

43. Vizcaíno, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

44. Davis, C. A. et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* **46**, D794–D801 (2018).

45. ENCODE Project Consortium An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

Acknowledgements

We thank T. Engleitner, R. Öllinger and R. Rad (TU Munich) for conducting the RNA sequencing run, mapping and normalization; the Quantitative Proteomics Unit at Goethe University for support with mass spectrometry measurements; and J. Riemer (University of Cologne) and R. Brandes (Goethe University, Frankfurt) for their critical feedback on the manuscript. This work was supported by the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme (Grant ERC StG 803565 to C.M.), the Deutsche Forschungsgemeinschaft

(German Research Foundation; Projects 390339347 (Emmy Noether Programme), 259130777 (SFB1177 Selective Autophagy), 456687919 (SFB1531 Stroma-vascular damage control) and 403765277 (mass spectrometer) to C.M., the Hessian Ministry for Arts and Sciences EnABLE Consortium and Fraunhofer High-Performance Center Innovative Therapeutics (TheraNova).

Author information

Author notes

1. These authors contributed equally: F. X. Reymond Sutandy, Ines Gößner

Authors and Affiliations

1. Institute of Biochemistry II, Faculty of Medicine, Goethe University Frankfurt, Frankfurt am Main, Germany

   F. X. Reymond Sutandy, Ines Gößner, Georg Tascher & Christian Münch

Contributions

F.X.R.S. and I.G. performed the experiments and data analysis. G.T. performed the mass spectrometry run and analysis of the DNAJA1 interactome. F.X.R.S., I.G. and C.M. interpreted the data and conceptualized the project. C.M. conceived and supervised this project and obtained funding. F.X.R.S. and C.M. wrote the paper with contributions from all authors.

Corresponding author

Correspondence to Christian Münch.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Navdeep Chandel, Xin Jie Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 The ATF4-ISR axis is not required for UPR^{mt} induction.

a, The experimental timeline of UPR^{mt} gene induction upon GTPP treatment. **b**–**d**, Time-resolved monitoring of relative transcript levels of (**b**) UPR^{mt} genes, (**c**) other mitochondrial proteins, and (**d**) ISR markers upon GTPP treatment measured with qPCR ($n = 4$ biological replicates). The line indicates the mean of transcript levels across different time points. **e**,**f**, Western blot images of (**e**) CHOP and (**f**) ATF4 in HeLa *CHOP* and *ATF4* knock-out cells, respectively in comparison to wild type (WT). "D" represents DMSO- and "G" represents GTPP-treated HeLa cells. Numbers indicate different clones used in the experiments. **g**, Barplots showing the mean of relative transcript levels of known ATF4 targets upon GTPP treatment in HeLa WT and *ATF4* knock-out cells measured with qPCR ($n = 3$ biological replicates). **h**, Barplots showing the mean of relative transcript levels of indicated UPR^{mt} genes upon GTPP treatment in HeLa *CHOP* and

ATF4 knock-out cells in comparison to WT measured with qPCR ($n = 3$ biological replicates). All P values are calculated with a two-tailed unpaired Student’s t-test and indicated in the figure. All error bars represent mean+SD. For gel source data, see Supplementary Fig. 1.

Source data

Extended Data Fig. 2 Time-resolved transcriptomic analyses reveal ROS requirement for UPR^{mt} induction.

a, Principal component analysis plot of transcriptomics from GTPP treated samples ($n = 3$ biological replicates). **b**, Heat map of increased (489) and decreased (383) transcripts within 3 h GTPP treatment. Transcript levels are represented as z-scores normalised for each gene (row z-score). **c**,**d**, Enrichment of GO BP terms from transcripts that are (**c**) increased or (**d**) decreased within 3 h GTPP treatment. Only GO terms with FDR < 0.1 are represented on the figures. **e**, Heat map representation of transcript levels of detected genes in the transcriptomic analyses that belong to GO BP: Transcriptional response to oxidative stress (GO:0036091). Transcript levels are represented as z-scores normalised for each gene (row z-score). **f**, Barplot showing the mean of mitochondrial ROS ($O_2^{\bullet-}$) levels of HeLa cells treated with GTPP measured on FACS using MitoSOX ($n = 3$ biological replicates). **g**,**h**, Barplots showing the mean of mitochondrial ROS ($O_2^{\bullet-}$) levels measured on FACS using MitoSOX (**g**) and relative transcript levels of indicated UPR^{mt} genes measured with qPCR (**h**) of HeLa cells upon co-treatments of GTPP with titrated concentrations of NAC ($n = 3$ biological replicates). **i**, **j**, Barplots showing the mean of mitochondrial ROS ($O_2^{\bullet-}$) levels measured on FACS using MitoSOX (**i**) and relative transcript levels of indicated UPR^{mt} genes measured with qPCR (**j**) of HeLa cells upon co-treatments of GTPP with titrated concentrations of GSH ($n = 3$ biological replicates). **k**,**l**, Barplots showing the mean of mitochondrial ROS ($O_2^{\bullet-}$) levels measured on FACS using MitoSOX (**k**) and relative transcript levels of indicated UPR^{mt} genes measured with qPCR (**l**) of HeLa cells upon co-treatments of GTPP with titrated concentrations of MnTBAP ($n = 3$ biological replicates). **m**, Barplot showing the mean of mitochondrial ROS

($O_2^{\bullet-}$) levels of HeLa cells upon co-treatments of GTPP with antimycin A or rotenone measured on FACS using MitoSOX (*n* = 3 biological replicates). All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD.

Source data

Extended Data Fig. 3 ROS are produced in mitochondria upon MMS.

a, Schematic illustration of HyPer7 fluorophore changes upon reaction with H_2O_2. Excitation of reduced and oxidized forms of the fluorophore are represented as light and dark green, respectively. **b**, Microscopy images of different constructs of HyPer7 targeted to the inter membrane space (IMS), the matrix (MTS) and untargeted (Cyto+Nuc) (*n* = 3). HyPer7 and mitotracker signals are shown in green and red, respectively. **c**–**e**, Representative microscopy images of HyPer7 targeted to the IMS (**c**), the matrix (**d**) or untargeted (**e**) of HeLa cells treated with GTPP for 3h (*n* = 5 biological replicates). Reduced (Ex405) and oxidized (Ex488) forms of HyPer7 are shown in green and red, respectively. **f**, Barplots showing the mean of H_2O_2 level measured with Hyper7 reporters targeted to the IMS, the matrix, and untargeted (cytosol+nucleus) upon 3h GTPP and different treatments as references (*n* = 5 biological replicates, 100 cells were analysed for each replicate). **g**, Schematic illustration for the mechanistic inhibition of VDAC1 by DIDS. **h**, Barplots depicting the mean of relative transcript levels of indicated UPR^{mt} genes of GTPP-treated HeLa cells upon co-treatments with DIDS (*n* = 4 biological replicates. **i**, Barplots showing the mean of H_2O_2 level measured with Hyper7 reporters targeted to the IMS, the matrix and untargeted (cytosol+nucleus) upon co-treatments of GTPP with DIDS (*n* = 5 biological replicates, >100 cells were analysed for each replicate). All scale bars indicate 50 µm. All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD.

Source data

Extended Data Fig. 4 DNAJA1 oxidation by MMS-induced ROS is an essential step in UPR^{mt} signalling.

a, Experimental scheme for redox proteomics upon GTPP treatment in HeLa cells. **b**, List of cysteines with significantly changed oxidation levels upon GTPP treatment ($P \leq 0.05$). **c**, Scheme of domain composition from different DNAJ family members. Amino acid (a.a.) lengths are indicated in brackets. ZFLR denotes the zinc finger like region. **d**, Barplots depicting the mean of relative transcript levels of *DNAJA2* and *DNAJB1* in HeLa *DNAJA2* and *DNAJB1* knock-down cells measured with qPCR, respectively ($n = 3$ biological replicates). **e**, Barplots depicting the mean of relative transcript levels of indicated UPR^{mt} genes of GTPP-treated HeLa cells upon knock-down of *DNAJA2* or *DNAJB1* measured with qPCR ($n = 3$ biological replicates). All P values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD.

Source data

Extended Data Fig. 5 DNAJA1 interacts with c-mtProt upon UPR^{mt} induction.

a, Enrichment map of GO Cellular Compartment (CC) terms of DNAJA1 interactors from the interaction proteomic analysis in HeLa cells. Circles represent GO terms with FDR<0.1. **b**, List of proteins with increased interactions to DNAJA1 upon GTPP treatment. **c**–**e**, Microscopy images of DNAJA1 (green), mitochondria (red), and nucleus (blue) upon GTPP treatment (**c**). Degrees of co-localization between DNAJA1 and mitochondria (Mitotracker) are represented both as a Pearson's correlation coefficient (**d**) and a Mander's overlap coefficient (**e**) with DNAJA1 as probe 1 and Mitotracker as probe 2 ($n = 3$ biological replicates). Boxplots represent the mean of Pearson's correlation coefficients (**d**), while barplot depicts the mean of Mander's overlap coefficients (**e**). Box indicates the interquartile range (IQR) and whiskers denote the 1.5 × IQR beyond the box. **f**. Changes in cell death upon GTPP treatment, measured by annexin V + propidium iodide staining and cell sorting. Viable cells are represented on the fourth quadrant (Q4), while non-viable cells on the second (Q2) and

third quadrants (Q3). Staurosporine (STS) was used as a positive control treatment. The mean of the percentage of non-viable cells is shown as a barplot in **g** ($n = 3$ biological replicates). Scale bar indicates 50 µm. All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD.

Source data

Extended Data Fig. 6 c-mtProt accumulate upon MMS.

a, Western blot images of HeLa *NRF1* knock-down efficiency in comparison to control siRNA (siControl). **b**,**c**, Barplots depicting the mean of relative transcript levels for several mitochondrial proteins (**b**) and indicated UPR^{mt} genes (**c**) in HeLa *NRF1* knock-down cells upon treatment with GTPP in comparison to siControl measured with qPCR ($n = 3$ biological replicates). **d**, Scheme of Halo-tagged protein reporter assay to monitor mitochondrial precursor protein import. HeLa cells stably expressing Halo-tagged mitochondrial proteins were incubated overnight with non-labelled Halo ligands to block the Halo tags of existing proteins. Newly synthesized proteins (non-blocked Halo tag) can be monitored with TMR-labelled Halo ligands. **e**, Representative microscopy images of cells stably expressing Halo-tagged ATP5A1. **f**,**g** Degrees of co-localization between Halo-tagged ATP5A1 and mitochondria (Mitotracker) are represented both as a Pearson's correlation coefficient (**f**) and a Mander's overlap coefficient (**g**). **h**, Representative microscopy images of cells stably expressing Halo-tagged GRPEL1. **i**, **j** Degrees of co-localization between Halo-tagged GRPEL1 and mitochondria (Mitotracker) are represented both as a Pearson's correlation coefficient (**i**) and a Mander's overlap coefficient (**j**). Both Halo-tagged protein quantifications were done with 3 biological replicates. For all Mander's overlap coefficient quantifications, Halo-tagged proteins were used as probe 1 and Mitotracker as probe 2. Boxplots (**g**,**i**) represent the mean of Pearson's correlation coefficients, while barplots (**h**,**j**) depict the mean of Mander's overlap coefficients. All boxes indicate the interquartile range (IQR) and whiskers denote the 1.5 × IQR beyond the box. All scale bars indicate 50 µm. All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD. For gel source data, see Supplementary Fig. 1.

Source data

Extended Data Fig. 7 HSF1 activation during UPR^{mt} is dependent on the translation of mitochondrial proteins.

a, Western blot images of mitochondrial chaperones in HeLa WT and *HSF1* knock-out cells. **b**,**c**, Barplots depicting the mean of relative transcript levels of (**b**) several UPR^{mt} markers and (**c**) an HSF1 target gene in HeLa *HSF1* knock-out cells upon treatment with GTPP in comparison to WT measured with qPCR ($n = 3$ biological replicates). **d**,**e**, Heat maps of relative transcript levels of indicated UPR^{mt} genes (mitochondrial chaperones and proteases) and other mitochondrial proteins upon (**d**) *HSF1* knock-out and (**e**) *NRF1* knock-down measured with qPCR ($n = 3$ biological replicates). **f**,**g**, Representative Western blot images of HSF1 in cytosolic (C) and nuclear (N) fractions of GTPP-treated HeLa cells upon co-treatment with (**f**) MnTBAP and (**g**) cycloheximide. **h**, Barplot depicting the mean of the nuclear to cytosolic HSF1 ratio (N/C) from the Western blots shown in **g** ($n = 3$ biological replicates). **i**, Representative Western blot images of HSF1 in cytosolic (C) and nuclear (N) fractions in HeLa *NRF1* knock-down cells in comparison to siControl. **j**, Barplot depicting the mean of the nuclear to cytosolic HSF1 ratio (N/C) from the Western blots shown in **i** ($n = 3$ biological replicates). **k**, Representative Western blot images of wild type FLAG tagged HSF1-HSP70 interactions upon GTPP treatment. **l**, Barplot depicting the mean of the HSP70 band intensities in comparison to HSF1 from the Western blots shown in **j** ($n = 3$ biological replicates). All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD. For gel source data, see Supplementary Figs. 1, 2.

Source data

Extended Data Fig. 8 HSF1 acts upstream of ATF5.

a, **b**, Barplots depicting the mean of relative transcript levels of (**a**) *ATF5* and (**b**) several UPR^{mt} markers in HeLa *ATF5* knock-down cells upon treatment with GTPP in comparison to siControl measured with qPCR ($n = 3$ biological replicates). **c**, Sequencing map showing the location of deleted

sequences on HeLa *ATF5* knock-out cells identified with Sanger sequencing. **d**,**e**, Barplots depicting the mean of relative transcript levels of (**d**) *ATF5* and (**e**) several UPR^{mt} markers in HeLa *ATF5* knock-out cells upon treatment with GTPP in comparison to WT measured with qPCR ($n = 3$ biological replicates). **f**, Time-resolved monitoring of relative transcript levels of *ATF5* upon 6 h of GTPP treatment ($n = 4$ biological replicates). The lines indicate the mean of transcript levels across different time points. **g**, HSF1 CHIP-seq data at the *ATF5* gene location. At the genome panel, boxes represent exons and strings represent introns. Read counts represent binding sites at the particular genome location. Binding sites were identified as read counts passing the irreproducible discovery rate threshold. **h**, HSF1 and ATF5 CHIP-seq data at the *HSPD1*, *HSPE1* and *HSPA9* gene locations in orange and black, respectively. All *P* values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD.

Source data

Extended Data Fig. 9 HSF1 increases cell survival upon MMS.

a, Representative microscopy images of HeLa cells transiently expressing MTS-GFP-tagged Abeta (green). Mitochondria were stained with Mitotracker (magenta). Scale bar indicates 20 µm. **b**, Barplots depicting the mean of relative transcript levels of indicated UPR^{mt} genes of HeLa cells upon overexpression of MTS-GFP tagged Abeta measured with qPCR at different time points post-transfection ($n = 2$ biological replicates). **c**, Degree of co-localization of MTS-EGFP and mitochondria (Mitrotracker) is represented as a Mander's overlap coefficient (M1) with MTS-EGFP as probe 1 and Mitotracker as probe 2 ($n = 5$ biological replicates). Barplot represents the mean of Mander's overlap coefficients. **d**–**f**, (**d**) Representative microscopy and (**e**) western blot images of WT and *HSF1* knock-out HeLa cells treated for 16 h with DMSO, 15 µM GTPP and 200 nM staurosporine (STS). Scale bar indicates 100 µm. Western blot images show the levels of apoptosis markers, including cleaved-PARP1 and cleaved-caspase 3 under the different treatments. Viability of cells under the treatments is calculated and represented as (**f**) barplot ($n = 6$ biological replicates). All *P* values are calculated with a two-tailed unpaired Student's

t-test and indicated in the figure. All error bars represent mean+SD. For gel source data, see Supplementary Fig. 1.

Source data

Extended Data Fig. 10 ROS and accumulation of c-mtProt are required as signals to induce the UPR^{mt}.

a, Measurement of mitochondrial ROS ($O_2^{\bullet-}$) levels on FACS with MitoSOX of HeLa cells treated with different conditions (left panel) and the mean of triplicate quantification as a barplot (right panel). **b**, Western blot images of mitochondrial proteins in their precursor (p) and mature (m) forms upon different treatments. **c**, Representative microscopy images of cells stably expressing Halo-tagged ATP5A1. **d**,**e** Degrees of co-localization between Halo-tagged ATP5A1 and mitochondria (Mitotracker) are represented both as a Pearson's correlation coefficient (**d**) and a Mander's overlap coefficient (**e**) (*n* = 3 biological replicates). **f**, Representative microscopy images of cells stably expressing Halo-tagged GRPEL1. **g**,**h** Degrees of co-localization between Halo-tagged GRPEL1 and mitochondria (Mitotracker) are represented both as a Pearson's correlation coefficient (**g**) and a Mander's overlap coefficient (**h**) (*n* = 3 biological replicates). For all Mander's overlap coefficient quantifications, Halo-tagged proteins were used as probe 1 and Mitotracker as probe 2. Boxplots (**d**,**g**) represent the mean of Pearson's correlation coefficients, while barplots (**e**,**h**) depict the mean of Mander's overlap coefficients. All boxes indicate the interquartile range (IQR) and whiskers denote the 1.5 × IQR beyond the box. Halo-tagged proteins and Mitotracker are shown in green and red, respectively. **i**, Gel image of mitochondrial insoluble and soluble fractions upon different treatments. Insoluble fractions are used as a measure for aggregate formation. **j**, Barplot showing the mean of mitochondrial ROS ($O_2^{\bullet-}$) levels measured on FACS using MitoSOX of HeLa cells treated with different mitochondrial stressors (*n* = 3 biological replicates). **k**, Western blot images of mitochondrial proteins in their precursor (p) and mature (m) forms upon treatment with different mitochondrial stressors. **l**, Barplots depicting the mean of relative transcript levels of indicated UPR^{mt} genes of HeLa cells upon treatment with different mitochondrial stressors measured with qPCR

($n = 3$ biological replicates). All scale bars indicate 50 µm. All P values are calculated with a two-tailed unpaired Student's t-test and indicated in the figure. All error bars represent mean+SD. For gel source data, see Supplementary Fig. 1.

Source data

Supplementary information

Supplementary Information

This file contains Supplementary Figs. 1–3 and Table 4.

Reporting Summary

Peer Review File

Supplementary Table 1

rlog normalized transcriptomics of 3-h GTPP treatment.

Supplementary Table 2

Modified peptides from redox proteomics of 3-h GTPP treatment.

Supplementary Table 3

DNAJA1 interaction proteomics upon GTPP treatment.

Source data

Source Data Fig. 1

Source Data Fig. 2

[**Source Data Fig. 3**]

[**Source Data Fig. 4**]

[**Source Data Extended Data Fig. 1**]

[**Source Data Extended Data Fig. 2**]

[**Source Data Extended Data Fig. 3**]

[**Source Data Extended Data Fig. 4**]

[**Source Data Extended Data Fig. 5**]

[**Source Data Extended Data Fig. 6**]

[**Source Data Extended Data Fig. 7**]

[**Source Data Extended Data Fig. 8**]

[**Source Data Extended Data Fig. 9**]

[**Source Data Extended Data Fig. 10**]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or

exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Sutandy, F.X.R., Gößner, I., Tascher, G. *et al.* A cytosolic surveillance mechanism activates the mitochondrial UPR. *Nature* **618**, 849–854 (2023). https://doi.org/10.1038/s41586-023-06142-0

- Received: 21 June 2022
- Accepted: 27 April 2023
- Published: 07 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06142-0

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06142-0

- Article
- Open Access
- Published: 14 June 2023

Genome expansion by a CRISPR trimmer-integrase

- Joy Y. Wang,
- Owen T. Tuck,
- Petr Skopintsev,
- Katarzyna M. Soczek,
- Gary Li,
- Basem Al-Shayeb,
- Julia Zhou &
- …
- Jennifer A. Doudna

Show authors

Nature **volume 618**, pages 855–861 (2023)

- 8171 Accesses
- 55 Altmetric
- Metrics details

Abstract

CRISPR–Cas adaptive immune systems capture DNA fragments from invading mobile genetic elements and integrate them into the host genome to provide a template for RNA-guided immunity[1]. CRISPR systems maintain genome integrity and avoid autoimmunity by distinguishing between self

and non-self, a process for which the CRISPR/Cas1–Cas2 integrase is necessary but not sufficient[2,3,4,5]. In some microorganisms, the Cas4 endonuclease assists CRISPR adaptation[6,7], but many CRISPR–Cas systems lack Cas4[8]. Here we show here that an elegant alternative pathway in a type I-E system uses an internal DnaQ-like exonuclease (DEDDh) to select and process DNA for integration using the protospacer adjacent motif (PAM). The natural Cas1–Cas2/exonuclease fusion (trimmer-integrase) catalyses coordinated DNA capture, trimming and integration. Five cryo-electron microscopy structures of the CRISPR trimmer-integrase, visualized both before and during DNA integration, show how asymmetric processing generates size-defined, PAM-containing substrates. Before genome integration, the PAM sequence is released by Cas1 and cleaved by the exonuclease, marking inserted DNA as self and preventing aberrant CRISPR targeting of the host. Together, these data support a model in which CRISPR systems lacking Cas4 use fused or recruited[9,10] exonucleases for faithful acquisition of new CRISPR immune sequences.

Main

Prokaryotes use CRISPR–Cas adaptive immune systems to create a sequential genetic record of infection[11]. Transcription and processing of CRISPR sequence arrays, which consist of short repeats and around 30 bp foreign DNA-derived spacers[1,2,12,13], yields mature CRISPR RNAs (crRNAs) that guide interference of matching genetic material, protecting the host against recorded sequences[14,15,16,17,18,19].

The Cas1–Cas2 integrase drives CRISPR array evolution by selecting and inserting new spacers[3,20,21]. $Cas1_4$–$Cas2_2$ is a heterohexameric complex that specifically recognizes DNA fragments (protospacers) containing an approximately 30 bp segment with short single-stranded 3′ overhangs[21,22,23]. In DNA-targeting CRISPR systems, protospacer selection requires a flanking 2–5 bp sequence known as the PAM, which is a key component used to distinguish self from non-self and evade autoimmunity. The PAM is selected during DNA capture, but is removed before host genome integration. Coordinated selection and removal of the PAM ensures Cas

interference modules target true invasive elements instead of the host CRISPR array[5,24,25].

Diverse mechanisms of PAM selection and removal underscore the importance of the PAM for maintaining both adaptive immunity and genome integrity during CRISPR sequence acquisition[8,9,10,26,27,28]. In CRISPR systems including type II-B, some type V, and type I-A, I-B, I-C, I-D and I-G, the Cas4 endonuclease performs PAM selection and processing[6,7,26,29,30]. However, around 40% of CRISPR subtypes lack Cas4[8]. In systems lacking Cas4 such as the type I-E system in the common laboratory *Escherichia coli* K12 strain, Cas1 contains a PAM-binding pocket that is believed to participate in protospacer precursor (prespacer) selection[9,23]. However, whether Cas1 cleaves the PAM in a similar manner to Cas4 or relies on host nucleases to perform this function remains unclear[9,23]. Recent in vitro studies identified host exonucleases that have the ability to aid Cas1–Cas2 in prespacer substrate trimming[9,10,31]. Standalone exonucleases such as the DnaQ-like exonuclease class DEDDh are widespread ancillary components that are present in every CRISPR–Cas type[8]. There are also type I-E systems containing a natural Cas2/DEDDh exonuclease fusion[8,28,32], further implying a functional link between exonucleases and the CRISPR integrase. These systems provide a model for studying coordination between host exonucleases and CRISPR integrases.

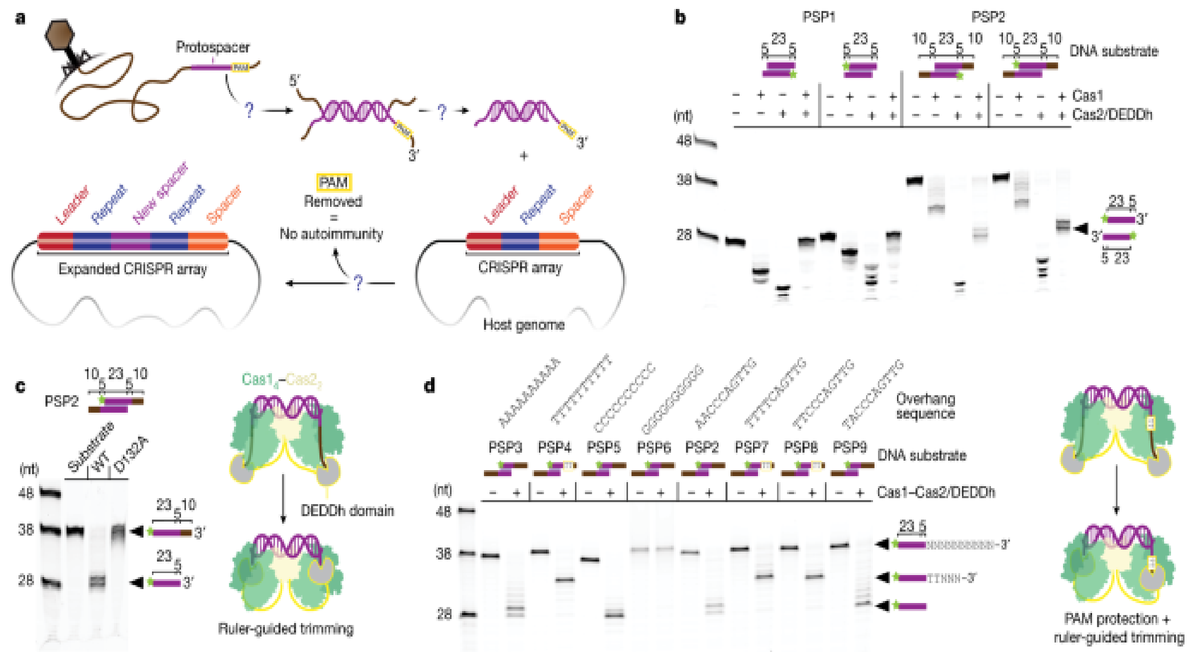
Here we reconstitute CRISPR sequence capture, processing and integration by a naturally occurring *Megasphaera* NM10-related Cas2 and DEDDh fusion protein (Cas2/DEDDh) in complex with Cas1 (Cas1–Cas2/DEDDh). We show that Cas1–Cas2/DEDDh preserves the PAM during prespacer processing and the first step of integration. The PAM is removed before completing full integration. The DEDDh active site, rather than Cas1[23], is responsible for both initial 3′ overhang trimming and PAM removal. This mechanism is distinct from that of Cas4, which cleaves the PAM endonucleolytically, suggesting a divergent role for host exonucleases in PAM processing[9]. The integrase regulates DEDDh exonuclease activity by a ruler-guided, gatekeeping mechanism that coordinates processing and defines the length of integrated DNA. Cryo-electron microscopy (cryo-EM) structures of Cas1–Cas2/DEDDh bound to prespacer DNA with or without the PAM show how Cas1–Cas2 recognizes the sequence and protects it from

DEDDh-mediated trimming. Conformational analysis of half-integration structures suggests that, once anchored into the CRISPR array, DNA bending engages the C-terminal region of Cas1, which in turn exposes the PAM for removal, enabling full integration. Our findings provide a general mechanism for exonuclease-assisted PAM processing and demonstrate that CRISPR systems evolved diverse mechanisms to ensure robust immunity against parasitic elements and avoid autoimmunity.

Cas1–Cas2/DEDDh substrate generation

CRISPR adaptation relies on the recognition, capture and processing of suitable DNA integration substrates from foreign sources (Fig. 1a). These integration substrates (prespacers) require nucleolytic processing to generate fragments of uniform length. To investigate the predicted exonuclease domain of Cas2/DEDDh, we expressed and purified Cas1 and Cas2/DEDDh from a type I-E *Megasphaera* NM10-related CRISPR system and tested DNA substrate processing in vitro (Fig. 1b). The size of spacers in the *Megasphaera* CRISPR array and preferences of the related I-E *E. coli* Cas1–Cas2 integrase suggest the preferred substrate is a 23 bp DNA duplex with 5 nucleotide single-stranded 3′ overhangs[22,23]. To test DNA processing, we assayed Cas1 and Cas2/DEDDh trimming activity using 5′ fluorophore-labelled prespacer substrates containing a 23 bp duplex region and extended single-stranded 3′ overhangs of varying lengths. Cas1 and Cas2/DEDDh each exhibit nuclease activity in isolation, yielding distinct products without apparent functional relevance (corresponding to partial cleavage or complete trimming of the 3′ end, respectively). Only the reconstituted Cas1–Cas2/DEDDh complex generates substrates equivalent in size to spacers in the host CRISPR array (Fig. 1b). Varying the substrate sizes showed that the integrase requires a 23 bp duplex for functional processing (Extended Data Fig. 1a,b). We next tested whether the DEDDh active site is responsible for processing activity by using a catalytically inactive DEDDh mutant (D132A). DNA cleavage assays indicated that the mutant complex does not process prespacers (Fig. 1c). Taken together, these data demonstrate that the complete Cas1–Cas2/DEDDh complex is necessary for prespacer processing and that the DEDDh active site provides the requisite nucleolytic activity.

Fig. 1: Cas1–Cas2/DEDDh processes prespacers to the correct size for integration and protects a TT PAM.



a, Open questions in CRISPR adaptation. **b**, Processing of fluorescently labelled prespacer substrates with a 23 bp duplex and different overhang lengths by Cas1–Cas2/DEDDh down to 5–6-nucleotide (nt) single-stranded 3′ overhangs. Prespacer 1 (PSP1), 23 bp duplex with 5-nucleotide 3′ overhangs; prespacer 2, 23 bp duplex with 15-nucleotide 3′ overhangs. The star indicates the 6-carboxyfluorescein label. **c**, WT and mutant Cas1–Cas2/DEDDh (DEDDh(D132A)) prespacer processing. **d**, Processing of substrates with variable 3′ overhangs and the model of observed PAM protection and ruler-guided trimming by the DEDDh domain. Gel source data are provided in Supplementary Fig. 1.

Time-course assays suggest similar processing efficiencies for prespacer substrates with varying overhang lengths (Extended Data Fig. 2a–d). Fluorescently labelled prespacers were incorporated into an integration target plasmid (pCRISPR) containing a shortened version of the natural *Megasphaera* CRISPR array. Kinetic analysis implies higher relative integration efficiency with the canonical substrate (23 bp duplex with 5-nucleotide single-stranded 3′ overhangs) compared with prespacers with extended overhangs. Reaction with the canonical substrate generated ligation products after 2 min, while prespacers with extended overhangs

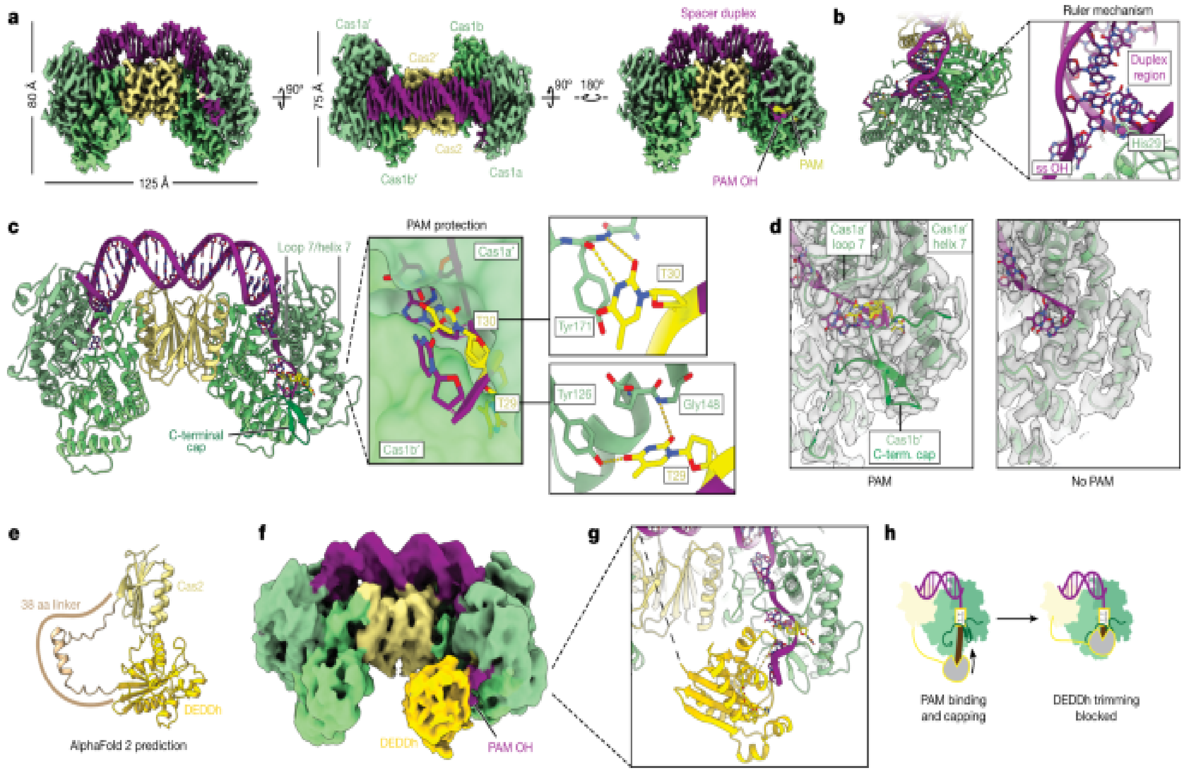
required 10 min for detection. Thus, Cas1–Cas2/DEDDh provides a molecular ruler against which DEDDh trims prespacers.

To determine the effect of the PAM, we varied the overhanging region, generating a small prespacer library against which the PAM could be inferred (Fig. 1d). We determined that Cas1–Cas2/DEDDh recognizes 5′-TT in the PAM position. In the absence of a TT PAM, DEDDh trims prespacer strands to the integration-competent size (28 nucleotides). However, the presence of a TT PAM in the correct position (nucleotide positions 29 and 30 relative to the 5′ end) results in partial trimming of the PAM-containing strand, precisely 3 nucleotides away from the PAM, yielding a 33-nucleotide product. We hypothesized that partial trimming was the result of sequestration by a PAM-binding pocket in Cas1[23].

PAM binding and prespacer processing

We next sought to elucidate the structural basis for prespacer processing and PAM protection. Cryo-EM was used to solve 3.1 Å and 2.9 Å resolution structures of Cas1–Cas2/DEDDh complexed with prespacer substrates with and without a TT PAM, respectively (Fig. 2a–d, Extended Data Table 1 and Extended Data Fig. 3). The *Megasphaera* Cas1–Cas2 retains the canonical heterohexameric architecture[33], with two Cas1 dimers (denoted Cas1 and Cas1′, a or b subunit) bridged by a central Cas2 dimer (Fig. 2a). A 23 bp duplex sits on the top of the complex, while 5-nucleotide single-stranded 3′ overhangs extend into the clefts formed at opposite Cas1 interfaces. The DEDDh domain could not be resolved in the PAM-absent dataset, presumably due to a high degree of conformational flexibility conferred by the 38 amino acid linker between DEDDh and Cas2, and because all 3′ ends are buried within the complex, protected from exonuclease activity (Fig. 2c). However, in the PAM-containing dataset, the DEDDh domain was resolved by iterative classification and three-dimensional refinement (Methods and Extended Data Fig. 4).

Fig. 2: Molecular detail of Cas1–Cas2/DEDDh during prespacer processing.



a, Orthogonal views of the final cryo-EM densities for Cas1–Cas2/DEDDh bound to a prespacer containing a phosphorothioated TT PAM (threshold, 0.200). **b**, The structure of PAM-bound Cas1–Cas2/DEDDh, depicting one of two His29 residues dictating duplex length. ss OH, single-stranded overhang. **c**, The structure viewed from the PAM side (left). Middle, surface depiction of the cleft between Cas1a′ and Cas1b′. Right, sequence-specific contacts made with each PAM thymine. **d**, Comparison of PAM and non-PAM densities with atomic models overlaid at a threshold of 0.200. Term., terminal. **e**, AlphaFold 2 prediction of the structure of Cas2/DEDDh. aa, amino acid. **f**, Side view of unsharpened cryo-EM density in which DEDDh was resolved (threshold, 0.033). **g**, Hybrid structure containing the DEDDh domain with detail at the PAM–DEDDh interface, with catalytic DEDDh residues shown. The black dashed line represents the unstructured linker between Cas2 and DEDDh domains. **h**, Model for PAM protection.

Cas1–Cas2/DEDDh dictates prespacer duplex length with an internal ruler that ensures that spacers are equivalent in length (Fig. 1b,c). Dual Cas1 His29 histidyl residues measure out a 23 bp DNA duplex by π-stacking with the terminal base pairs of the double-stranded region, clasping the prespacer strands and marking the start of the single-stranded 3′ overhang (Fig. 2b).

Biochemical processing assays demonstrate that only the 23 bp duplex is both tolerated and trimmed to the integration-competent length (Extended Data Fig. 1a,b).

Processing experiments suggest that the PAM is protected initially from DEDDh-mediated prespacer processing (Fig. 1b–d). To understand how the integrase sequesters the PAM, we performed cryo-EM analysis of the integrase complex bound to a PAM-containing prespacer with a phosphorothioate backbone modification at the predicted site of DEDDh activity, with the intention of stalling processing. Sequence-specific interactions with loop 7 (Asn162 to Asp179) and helix 7 (Met139 to Tyr161) in the Cas1a′ subunit in the resultant density rationalize PAM recognition (Fig. 2c). The first PAM thymine is buried in a pocket in Cas1a′, where hydrogen bonds formed with Tyr126 and Gly148 may enhance binding affinity (Fig. 2c (top right)). The second PAM thymine π-stacks with Tyr171, which positions T30 to hydrogen bond with the Tyr171 backbone amide nitrogen (Fig. 2c (bottom right)). When the PAM is absent, the substrates are fully trimmed, underscoring the necessity of sequence-specific interactions for asymmetric trimming and PAM protection (Fig. 1d). Moreover, a β-hairpin in the C-terminal region of the Cas1b′ 'caps' sequestered nucleotides. Loop 7, helix 7 and the C-terminal cap are absent in the PAM-deficient density, suggesting that these structural motifs participate in PAM protection (Fig. 2d).

The DEDDh domain was not visible in the initial PAM-containing structure, raising the question of how the integrase performs ruler-guided trimming of sequestered nucleotides. To resolve the DEDDh domain, we iteratively classified and refined PAM-containing particles and found a density corresponding to DEDDh in a small subset of the total ensemble (Fig. 2f and Extended Data Fig. 4). Key features include a large protrusion only visible on the PAM side of the complex and an extended density attributable to additional phosphorothioate-containing nucleotides 3′ to the PAM (Fig. 2f). As the protuberance was low resolution, the model predicted by AlphaFold 2 for the DEDDh domain was docked into the density (Fig. 2f,g). The resulting hybrid model illustrates dynamics of DEDDh trimming and PAM protection. Catalytic DEDDh residues are poised to exonucleolytically cleave the overhang, but the PAM-binding pocket and the C-terminal loop occlude DEDDh procession, blocking cleavage of the PAM and 2–3

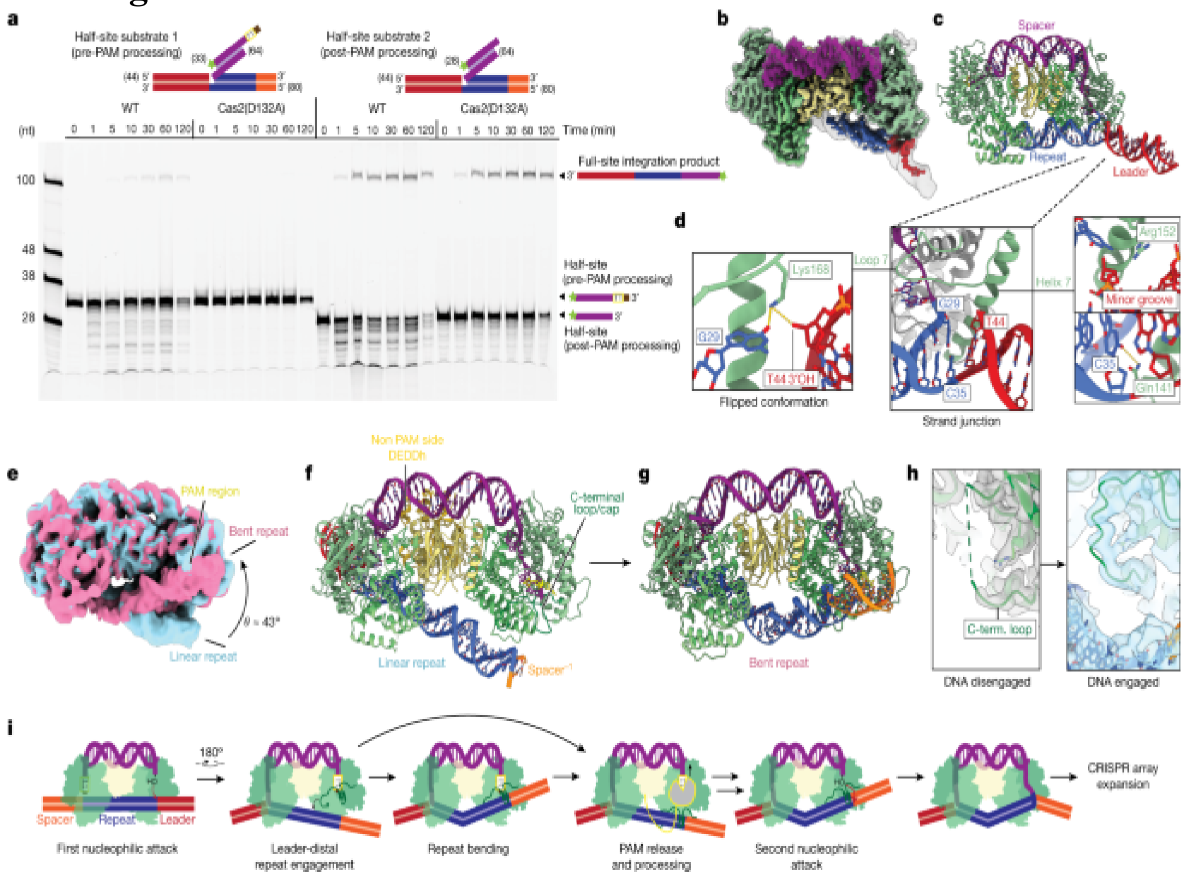
additional nucleotides (Fig. 2h). Despite the high local concentration of non-specific exonuclease relative to the substrate, this process is precise, in concordance with biochemical evidence (Figs. 1b and 2g). A natural consequence of protection is that the PAM must be cleaved downstream.

PAM trimming after half-integration

Although there is evidence for PAM protection during prespacer processing, the PAM must be removed before insertion into the CRISPR array to avoid autoimmunity. We analysed Cas1–Cas2/DEDDh processing of DNA substrates designed to mimic intermediates of integration into the CRISPR array to determine the mechanism of PAM removal and resolve dynamics of the complex at the integration target site[7,9]. Two substrates that mimic probable half-integration intermediates—the pre-PAM processing intermediate and the post PAM-processing intermediate (Fig. 3a; half-site substrates 1 and 2, respectively)—were synthesized and assayed in reactions with wild-type (WT) and catalytically inactivated DEDDh complexes. Reaction with half-site substrate 1, which contains the unprocessed PAM, resulted in a 100-nucleotide band corresponding to full-site integration (Fig. 3a). The same band was absent when DEDDh was catalytically inactive. Reactions with PAM-deficient half-site substrate 2 yielded the 100-nucleotide full-site integration product for both the WT and dead DEDDh complexes. These data suggest that, in the WT reaction, the PAM is fully removed before full-site integration. The lower intensity of the full-site integration product generated from half-site substrate 1 compared to that of half-site substrate 2 may be a result of inefficient PAM removal, also observed in kinetics assays (Extended Data Fig. 2a,b). The absence of the integrated product strand in the catalytically inactivated DEDDh condition suggests the DEDDh active site executes PAM processing. Notably, PAM processing is necessary for full-site integration, and the integrase generates a precisely defined insertion product size. Thus, non-specific exonuclease activity generates a ladder of ssDNA overhang fragments in the PAM-containing substrate strand, but only one of these fragment sizes is compatible with full-site integration. This single-nucleotide precision is a result of the Cas1–Cas2 ruler, which simultaneously defines the spacer size and acts as a gate that prevents PAM insertion into the CRISPR array (Fig. 3a). As Cas1-mediated PAM protection was observed during prespacer

processing, it is reasonable to assume that Cas1 releases the PAM for DEDDh trimming while engaged on the CRISPR array.

Fig. 3: Biochemical and structural analysis of Cas1–Cas2/DEDDh PAM processing.



a, Integration reactions by WT and mutant Cas1–Cas2/DEDDh (DEDDh(D132A)) half-site integration intermediates. Gel source data are provided in Supplementary Fig. 1. **b**, Unsharpened (grey transparent; threshold, 0.05) and sharp (colour; threshold, 0.19) cryo-EM densities for Cas1–Cas2/DEDDh bound to the PAM-phosphorothioated half-site integration analogue. **c**, The structure of the initial half-site complex. **d**, Details of the junction of leader, repeat and spacer (middle). Left, a second, flipped conformation of G29. Right, helix 7 interactions. **e**, Unsharpened maps of two repeat DNA conformations, linear (blue) and bent (red), both with a threshold of 0.05. **f**, Structure with linear, extended repeat DNA. **g**, Structure of bent repeat DNA. **h**, Comparison of C-terminal loop ordering in the prespacer (left) and linear repeat (right) structures. **i**, Model for PAM gatekeeping facilitated by the C-terminal loop/cap.

Evidence for DEDDh involvement in both prespacer processing and PAM cleavage led us to examine which molecular cues prompt Cas1 to relinquish the PAM for digestion. Aiming to visualize DEDDh trimming and conformational changes in Cas1, we used cryo-EM to characterize Cas1–Cas2/DEDDh in complex with a DNA half-site analogue containing phosphorothioate linkages at the PAM positions (Fig. 3b and Extended Data Fig. 5a–c). Neither the initial 3.1 Å density nor any heterogeneous states detected during cryo-EM data processing had density corresponding to the DEDDh domain on the PAM side of the complex (Fig. 3b). However, DEDDh was observed on the non-PAM side during DNA conformational analysis. We speculate that, in agreement with biochemical data, only the DEDDh domain can trim the PAM, and PAM trimming activity is required for full integration (Fig. 3a). Furthermore, the DEDDh PAM-trimming state may be transient. Phosphorothioate modifications only partially protected the PAM, which may disfavour resolution of the active DEDDh domain at the half-site (Extended Data Fig. 6).

The 3.1 Å half-site structure reveals interactions at the first integration strand junction (Fig. 3b–d and Extended Data Table 1). After the initial nucleophilic attack, Cas1–Cas2 induces bending of the leader-repeat target DNA[34]. This deformation originates at the nick site, positioning repeat DNA between the leader-distal and -proximal Cas1 active sites. Inspection of the strand junction (Fig. 3d (middle)) reveals interactions with the first base pair of the CRISPR repeat dictated by loop 7 and helix 7 of Cas1a′. The spacer-ligated first CRISPR repeat guanine G29 was found in two approximately equivalent conformations. In the first conformation, G29 forms a canonical base pair with C35 (Fig. 3d (middle)). In the second conformation, G29 flips upwards, making a specific contact with loop 7 lysine Lys168. The lysine also contacts the 3′ hydroxyl of the leader (Fig. 3d (left)). Gln141, which sits at the base of helix 7, also makes a nucleobase-specific contact with C35, the first bottom repeat nucleotide. Helix 7 is well positioned to insert into the minor groove of the leader DNA, but no nucleobase-specific contacts were obvious (Fig. 3d (right)). Specific contacts with CRISPR array nucleotides probably have a functional role in targeting, as previously observed[20,21,35].

The 3.1 Å half-site complex reconstruction contains density corresponding only to the leader-proximal region of the CRISPR repeat (Fig. 3b,c). To probe the dynamics at the leader-distal region, where PAM processing and

subsequent full integration occur, we performed three-dimensional variability analysis (3DVA) of the particle set (Extended Data Fig. 5a,d,e)[36]. 3DVA revealed heterogeneity in the location of the repeat/spacer end, with the CRISPR repeat DNA oscillating between linear and bent conformations (Fig. 3e and Supplementary Video 1). The DEDDh domain was visible only in the linear conformation (Supplementary Video 2). Isolation and refinement of particle clusters representing maxima of the reaction coordinate gave linear and bent reconstructions at resolutions of 4.1 Å and 3.9 Å, respectively (Fig. 3e–g).

In the linear structure, a Cas1b′ C-terminal loop (Leu279 to Ser293) rich in charged residues is positioned near to the major groove adjacent to the second integration target site (Fig. 3f). The corresponding density is absent in the 3.1 Å half-site and 2.9 Å prespacer-bound structures, indicating that this loop participates in engagement with the CRISPR repeat on the PAM side (Figs. 2c and 3g,h). A C-terminal cap, which follows the C-terminal loop, protects the PAM and adjacent nucleotides from trimming by DEDDh (Fig. 2h). Notably, the DEDDh domain was visible in the linear structure, but on the non-PAM side of the complex, where no overhang trimming occurs (Fig. 3f). The exonuclease sits in a cavity formed by the interface of the CRISPR repeat DNA, Cas1a/b′ and Cas2, where it appears to contact the repeat DNA backbone and the N terminus of Cas1b (Fig. 2f and Extended Data Fig. 7). These interactions may beneficially constrain bending of the second integration target site or prevent *trans* activity.

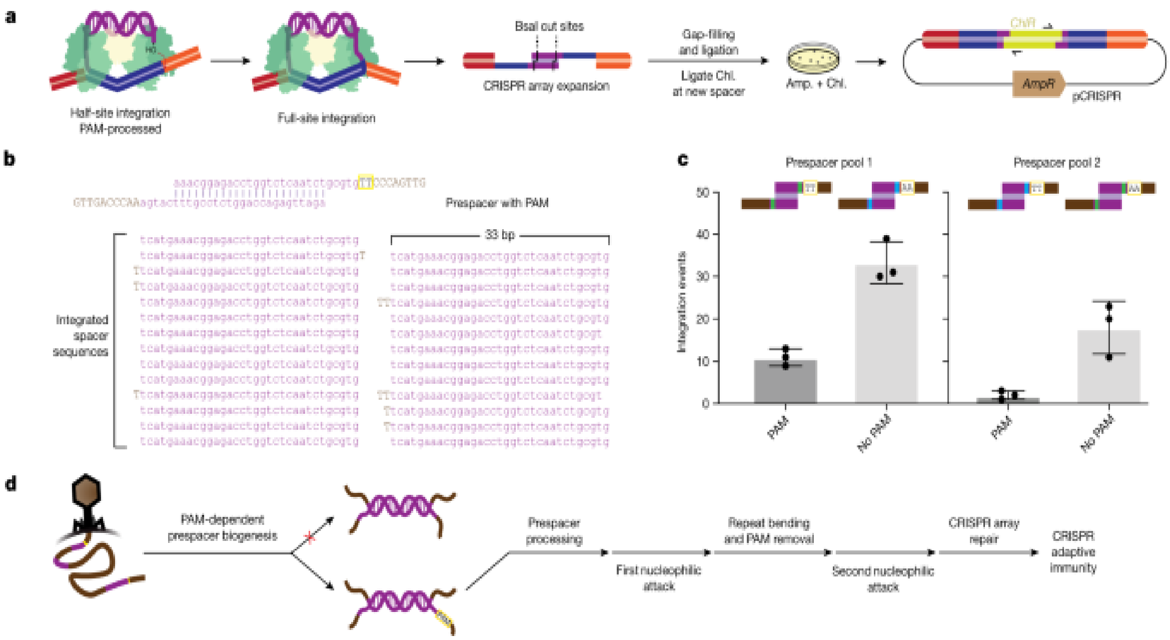
The bent structure features a pronounced kink in the centre of the repeat region (Fig. 3g). Disruption of a single A–T base pair and DNA unwinding appear to accommodate the strain induced by this pitch, although this assignment was made with low confidence owing to the low local resolution at the bending site. Bending in the centre of the CRISPR repeat symmetrizes the integration complex and draws the repeat/spacer junction towards the Cas1a′ active site. Although PAM nucleotides are still present, probably due to their cleavage-blocking phosphorothioate modifications, the overhang nucleotides typically sequestered during ruler-guided trimming are absent, suggesting that 3′ trimming occurs in an intermediate step between the linear and bent states. Only the first three nucleotides under the C-terminal loop of Cas1b′ could be assigned, and the C-terminal cap density was largely unstructured. These observations, combined with the proximity of the

second integration target site in the bent structure, suggest that engagement of the CRISPR DNA by the C-terminal loop induces uncapping of previously sequestered PAM nucleotides, followed by DEDDh-mediated or host-exonuclease-mediated trimming of the exposed 3′ end to the ruler-defined length. Once the PAM is trimmed, full integration occurs (Fig. 3a). Structural and biochemical analyses imply a general mechanism of sequential PAM protection and cleavage, or PAM gatekeeping, which ensures that PAM-deficient protospacers integrated into the CRISPR array are marked as self and are equal in size (Fig. 3i).

CRISPR array integration reconstitution

To determine how PAM sequence recognition and gated removal ensures accurate DNA integration, we reconstituted CRISPR substrate integration in vitro. An unprocessed prespacer (23 bp duplex with 15 nucleotide single-stranded overhangs) containing the TT PAM was combined with Cas1–Cas2/DEDDh and pCRISPR. Prespacer substrates encoded BsaI restriction sites in the duplex region to enable insertion of a chloramphenicol-resistance gene. After transformation, only pCRISPR with full prespacer integration confers survival in a double selection assay[37] (Fig. 4a). Fully integrated sequences provide additional evidence that Cas1–Cas2/DEDDh completes both prespacer processing and integration into the CRISPR array (Fig. 4b). The complex is specific for the CRISPR array and all integration events occur at repeat borders (Extended Data Fig. 8b). However, integration occurs at all three repeats present in the array, without specificity for the leader-proximal repeat, as seen in many CRISPR systems in vivo[1,38] and consistent with structural data (Fig. 3d). Excess 3′ overhangs are trimmed to within 1–2 nucleotides of the expected length and the PAM is absent in all integrated sequences (Fig. 4b), in agreement with evidence at the level of the half-site (Fig. 3a,i). Reconstitution experiments provide complementary evidence for an alternative mechanism for PAM processing compared with Cas4, which uses a sequence-specific mechanism to cleave the PAM endonucleolytically[7,29].

Fig. 4: In vitro reconstitution of Cas1–Cas2/DEDDh-mediated full-site integration.



a, Schematic of in vitro reconstitution of full-site integration. Amp., ampicillin; chl., chloramphenicol. **b**, Fully integrated spacer sequences from a prespacer containing the PAM. Lowercase bases match the prespacer sequence. Uppercase bases are additional, untrimmed nucleotides. **c**, The number of integration events arising from an equimolar pool of prespacers with or without the PAM. Prespacers are distinguished by internal barcodes, which were swapped to remove sequence bias. Data are mean ± s.d. of three independent biological replicates ($n = 95$ colonies). **d**, Proposed general timeline of CRISPR adaptation.

Source data

Although it was hypothesized that delayed PAM trimming aids the complex in orienting the prespacer for integration[7,9], no orientation bias was observed in vitro (Extended Data Fig. 8c). Although Cas1–Cas2/DEDDh alone is able to distinguish between the PAM and non-PAM sides of the prespacer (Figs. 1d and 2c), it appears that the complex alone cannot discern the leader- and spacer-side of the repeat, consistent with cryo-EM results of the half-site intermediate, which show no sequence specificity for the leader (Fig. 3d). We suspect that the complex requires additional host factors to correctly orient spacers in vivo. In *E. coli*, integration host factor (IHF) directs the first nucleophilic attack to the leader-side of the repeat through specific contacts with the leader sequence[34,39]. Superimposition of the half-site

structure and a *Megasphaera* IHF orthologue onto a structure of the complete IHF-containing integration holo complex further implicates the participation of a directing host factor (Extended Data Fig. 9). In vivo, the system may have higher specificity for the leading integration target site and use delayed PAM processing as the basis for determining the orientation of integration, as is the case in other CRISPR systems[7,9,10].

To assess the effect of the PAM on integration efficiency, an equimolar mixture of PAM-deficient and PAM-containing prespacers, each containing a pair of identifying internal barcodes, was tested for full-site integration. Notably, out of 95 sequenced colonies, we observed significant enrichment (around threefold) of integration events from the PAM-deficient prespacer (Fig. 4c). To account for biases resulting from the internal barcode sequences, we generated a second prespacer pool, in which barcode pairs were swapped. The second pool also exhibited a significant preference for the PAM-deficient prespacer (Fig. 4c). Lower integration efficiency from the PAM-containing prespacer in vitro may stem from additional steps that are required for PAM removal (Fig. 3i). Moreover, PAM removal is observed after full-site integration, and all spacer sequences are selected according to PAM presence (Fig. 4b). The reduced apparent efficiency of PAM-containing prespacer insertion in vitro therefore suggests PAM recognition in vivo occurs upstream, during the biogenesis of substrates bound for CRISPR adaptation (Fig. 4d).

Integration reconstitution experiments with pooled prespacers suggest that the integrase may select substrates before prespacer processing, ensuring PAM presence (Fig. 4c). We were interested in whether the integrase demonstrates similar stringency for another substrate feature—the canonical prespacer duplex. Accordingly, we tested Cas1–Cas2/DEDDh processing after stepwise addition of the PAM complementary strand. After incubation of Cas1–Cas2/DEDDh with a single-stranded PAM-containing strand, the labelled PAM-deficient strand was added. In all of the reactions, even for the prespacer strands with the tolerated 23-nucleotide complementarity region (Fig. 1b), non-specific processing of the labelled strand occurs (Extended Data Fig. 8d–f). Thus, Cas1–Cas2/DEDDh probably performs ruler-guided trimming when the 23 bp prespacer is preduplexed and not after delayed addition or search for the complementary strand. The strict requirement for

substrate size, strandedness and PAM presence has implications for open questions in CRISPR substrate biogenesis.

Discussion

Efficient CRISPR adaptive immunity requires coordination between the CRISPR integrase and host nucleases[9,10]. In this study, we describe mechanisms of prespacer processing and integration in a naturally occurring Cas1–Cas2/DEDDh complex. The trimmer-integrase uses an alternative PAM-processing mechanism compared with the well-studied Cas4 endonuclease[7,30]. Previously, it was unclear how systems lacking Cas4 process and integrate substrates. Our data suggest that one evolutionary solution to the problem of selecting, protecting and then removing the PAM is to use Cas1 rather than an accessory protein for initial PAM protection. Sequestration of defined prespacer sizes through substrate gatekeeping ensures that the PAM is present and that its cognate spacer is functional (Fig. 1d). Once the PAM-containing prespacer is anchored to the host CRISPR array, the PAM is released by the Cas1 gate and is promptly removed. We provide a mechanism explaining which structural cues lead to PAM uncapping and removal (Fig. 3i). Binding and bending of leader-distal repeat DNA may lead to disengagement of the C-terminal cap, which covers and protects nucleotides. DEDDh completely digests the released PAM, generating substrates of the correct size and positioned for second nucleophilic attack. This sequence of events ensures that the PAM side integrates second. Bending and unwinding may also aid in the melting of the repeat strand, which is required for resolution of the post-synaptic complex and concomitant repeat duplication[40]. Although the high effective local concentration of exonuclease with respect to the bound prespacer conferred by the Cas2/DEDDh fusion in this system probably improves efficiency, we imagine that host exonucleases can function similarly in *trans*. Cas1–Cas2/DEDDh serves as a general model describing the role of accessory exonucleases, including those that are not fused to the nuclease, in the diverse CRISPR systems lacking Cas4 (comparisons to *E. coli* Cas1–Cas2 are shown in Extended Data Fig. 10).

Recent studies indicated that accessory proteins can coordinate with Cas1–Cas2 to process prespacers. DnaQ and ExoT have previously been shown to

process PAM-containing substrates asymmetrically when provided in concert with Cas1–Cas2 in vitro[9,10], establishing the directionality of integration. Recent in vivo research demonstrated that other accessory exonucleases can substitute for DnaQ and ExoT activities to carry out prespacer processing[31]. These findings suggest that Cas1–Cas2 can flexibly coordinate with various accessory proteins. As a general model, Cas1–Cas2/DEDDh provides insights into the elegant mechanism by which non-specific processing enzymes and Cas1–Cas2 preserve the self versus non-self distinction. These findings advance our understanding of how prespacers are processed and selected for spacer acquisition. We anticipate that our results will be applicable to CRISPR-based technologies that seek to repurpose Cas1–Cas2 for molecular recording and information storage, applications challenged both by reliance on host factors such as exonucleases and by uncertainty in prespacer selection[41,42,43].

Although this report represents an advance in our understanding of downstream steps, the upstream biogenesis of CRISPR substrates remains unclear. Unexpectedly, experimental data suggest lower integration efficiency from PAM-containing prespacers and a preference for preformation of the canonical duplex[22]. These results weaken the ‘complement search’ model for prespacer biogenesis, which suggests single-stranded DNA derived from foreign sources are captured independently by the integrase complex. Alternatively, Cas1–Cas2 may recognize PAM-containing prespacer-like motifs as DNA reanneals behind repair complexes implicated in CRISPR adaptation such as RecBCD and AddAB, as was recently suggested[33]. The precise mechanistic details of this proposal are unclear. Future experiments might use the compact trimmer-integrase presented here to investigate open questions in CRISPR substrate biogenesis and achieve total in vitro reconstitution of naive CRISPR adaptation.

Methods

Plasmid construction and DNA substrate preparation

To make the target integration plasmid pCRISPR, the leader and the first three repeats and spacers of the CRISPR array were ordered as two DNA

fragments, which were amplified by PCR and inserted into the pUC19 backbone by Gibson assembly. DNA oligos used in this study were ordered from Integrated DNA Technologies. Prespacers and the half-site substrates were formed by heating at 95 °C for 5 min and slow cooling to room temperature in HEPES hybridization buffer (20 mM HEPES, pH 7.5, 25 mM KCl and 10 mM $MgCl_2$). For the half-site substrate, hybridization was performed with a 1.5-fold excess of the two shortest strands and a 1.25-fold excess of the second-largest strand and purified on an 8% native PAGE gel. Sequences of cloning primers and DNA substrates are shown in Supplementary Table 2.

Cloning, expression and purification

The *Megasphaera* NM10-related *Cas1* and *Cas2-DEDDh* genes were codon-optimized for *E. coli* expression, ordered as G-blocks, PCR-amplified and cloned separately into a pET-based expression vector with an N-terminal 10×His-MBP-TEV tag. After transformation into chemically competent Rosetta cells, cells were grown to an optical density at 600 nm of around 0.6 and induced overnight at 16 °C with 0.5 mM isopropyl-β-d-thiogalactopyranoside. Cells were collected and resuspended in lysis buffer (20 mM HEPES, pH 7.5, 500 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, 1 mM Tris (2-carboxyethyl)phosphine (TCEP), Complete EDTA-free protease inhibitor (Roche), 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and 10% glycerol). After lysis by sonication and clarification of the lysate by centrifugation, the supernatant was incubated with Ni-NTA resin (Qiagen). The resin was washed with wash buffer (20 mM HEPES, pH 7.5, 500 mM NaCl, 10 mM imidazole, 1 mM TCEP and 5% glycerol) and the protein was eluted with wash buffer supplemented with 300 mM imidazole. After overnight digestion with TEV protease, the salt concentration was diluted to 300 mM NaCl using ion-exchange buffer A (20 mM HEPES, pH 7.5, 1 mM TCEP and 5% glycerol) and run through a tandem MBPTrap column (GE Healthcare) and HiTrap heparin HP column (GE Healthcare) to remove the MBP and bind the protein onto the heparin column. The protein was eluted with a gradient from 300 mM to 1 M KCl, concentrated and purified on the Superdex 200 (16/60) column with storage buffer (20 mM HEPES, pH 7.5, 500 mM KCl, 1 mM TCEP and 5% glycerol). The same purification protocol was used for Cas1 and Cas2/DEDDh (WT and D132A

mutant). The sequences of the proteins are provided in Supplementary Table 1.

Processing assays

Processing assays were conducted in integration buffer (20 mM HEPES, pH 7.5, 125 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Nonidet P-40 and 10% DMSO). Cas1 (4 µM) and Cas2/DEDDh (2 µM) were precomplexed for 30 min at 4 °C before addition of fluorescent DNA substrate (312.5 nM) and reacting for 2 h at 37 °C. The reaction was quenched by addition of 2 vol quench buffer (95% formamide, 30 mM EDTA, 0.2% SDS and 400 µg ml^{-1} heparin) and heating at 95 °C for 4 min, before analysis on a 14% urea–PAGE gel. Reactions were visualized using the Typhoon FLA gel imaging scanner and quantification of intensities was performed using ImageQuantTL (v.8.2). The percentage processing activity was quantified as the ratio of the final product band intensity to the total intensity of all bands in the lane.

Cryo-EM data acquisition

Cas1–Cas2/DEDDh DNA complexes were formed by mixing 50 µM Cas1, 50 µM Cas2/DEDDh, and 12.5 µM prespacer or half-site DNA, and dialysing for 2 h using a Slide-A-Lyzer MINI Dialysis Device at room temperature. The complex was concentrated to varying concentrations of Cas1–Cas2/DEDDh (Extended Data Table 1) and purified over the Superose 6 Increase 10/300 GL column. The samples were frozen using the FEI Vitrobot Mark IV, cooled to 8 °C at 100% humidity. Depending on the sample (Supplementary Table 1), either carbon 2/2 300 mesh C-flat grids (Electron Microscopy Sciences CF-223C-50) or 1.2/1.3 300 mesh UltrAuFoil gold grids (Electron Microscopy Sciences, Q350AR13A) were glow discharged at 15 mA for 25 s using PELCO easyGLOW. In all cases, a total volume of 4 µl sample was applied to the grid and immediately blotted for 5 s with a blot force of 8 units. Micrographs were collected on the Talos Arctica operated at 200 kV and ×36,000 magnification (1.115 Å pixel size), in the super-resolution setting of the K3 Direct Electron Detector. Cryo-EM data were collected using SerialEM (v.3.8.7). Images were obtained in a series of exposures generated by the microscope stage and beam shifts.

Cryo-EM data processing

All datasets were collected with varied tilt angle, number of videos and defocus range (Supplementary Table 1 and Extended Data Figs. 3–5). Data processing was performed in cryoSPARC (v.3.2.0, v.3.3.1 and v.4.1.1)[44]. Videos were corrected for beam-induced motion using patch motion correction, and contrast transfer function parameters were calculated using patch CTF.

The PAM-deficient prespacer-bound Cas1–Cas2/DEDDh map was obtained through an iterative process. In the first round, 569 particles were picked manually from 37 micrographs and submitted for Topaz training[45]. The resulting Topaz model was used to pick particles from the micrographs, and a total of 460,631 particles was extracted with a bin factor of 2, and applied to 2D classification. After selecting the best classes, 410,757 particles were used for ab initio reconstruction and subsequent heterogenous refinement, with three classes. All of the particles were used for non-uniform map refinement[46], and an initial complex map was obtained. After 2D classification of particles from the initial non-uniform refinement model, 38,342 particles from the classes with isotropic orientations were selected and processed for the second round of Topaz training. A new Topaz model was used with a total of 956 curated micrographs, and the entire process was repeated twice with particles from the best heterogeneous refinement class for subsequent non-uniform refinement and Topaz training. The final map with the best electron density for the PAM-deficient prespacer bound Cas1–Cas2/DEDDh complex was obtained from 461,266 particles and was refined with non-uniform refinement to 3.1 Å.

For the PAM-containing prespacer-bound Cas1–Cas2/DEDDh, a single round of Topaz training was applied. After the initial exposures curation, which yielded 591 best-quality micrographs, 6,302 particles were manually picked and processed for the Topaz training job. The Topaz model was applied to an expanded set of 1,184 curated micrographs, and resulted in extraction of 3,101,776 particles. After ab initio reconstruction and heterogenous refinement of the particles, with three classes, the 1,420,721-particle set constituting the best class were processed with non-uniform

refinement. As a result, a 2.9 Å density for PAM-containing prespacer bound Cas1–Cas2 complex was obtained.

For resolving the DEDDh density in the latter dataset, the ab initio class particles used for the latter density reconstruction, 1,331,357 in total, were applied to a 2D classification job, and 228,220 particles were selected in classes with apparent DEDDh density. After ab initio refinement with three classes, particles from the best class were processed for another round of 2D classification, and 109,912 particles with more pronounced DEDDh density were selected, and re-extraction was performed with a 320 pixel box size (in all other cases, 480 pixel boxes were used for the extraction jobs). As a result of the final 2D classification round, 49,560 particles with the best DEDDh density were selected, re-extracted with standard box dimensions and processed for ab initio refinement, with one class and non-uniform refinement. As a result, a 3.5 Å complex map with the DEDDh exonuclease density was obtained, with a total of 49,383 particles used for reconstruction.

For half-site DNA-bound Cas1–Cas2/DEDDh, the Topaz model from the PAM-containing prespacer was applied to 2,810 micrographs selected after manual curation. The 2,448,888 resultant particles were subdivided using 2D classification, and the 25 best classes were selected, resulting in 1,836,610 particles. These particles were processed for ab initio reconstruction with three classes. The best class containing 1,048,353 particles was refined using non-uniform refinement to yield to the 3.1 Å half-site map.

To observe DNA dynamics in the Cas1–Cas2/DEDDh half-integration complex, we performed 3DVA[36] on a subset of particles selected and refined from 2D classification with DNA visible on the leader-distal side of the complex (1,048,353 particles). The filter resolution was 6 Å and the number of modes was 3. To generate Supplementary Video 1, the 3DVA output mode was set to simple and 20 frames, then UCSF ChimeraX was used to generate a vseries. Next, the 3DVA output mode was set to cluster and the number of clusters was set to 20. Each resulting cluster was individually inspected, and two clusters representing maxima of DNA motion along the pitch axis were chosen. The linear structure was derived from 32,722 particles and was processed for non-uniform refinement to give the final 4.1 Å map. The bent structure resulting from initial 3DVA clustering was improved by repetition of the 3DVA workflow with the complete particle set

obtained by Topaz picking, then selection and non-uniform refinement of the cluster representing leader-distal DNA in the most bent conformation (53,545 particles total), yielding the final 3.9 Å map.

Model building and refinement

The initial models of the Cas1 and Cas2/DEDDh were obtained using the AlphaFold 2 program[47]. To build the model of Cas1–Cas2/DEDDh bound to a prespacer with TT PAM complex, the predicted Cas1 and Cas2 monomers were docked independently into the corresponding map with the fitmap tool in UCSF ChimeraX (v.1.2.5)[48]. The DNA models were built de novo. The complex model was refined using rounds of real-space refinement and rigid body fit tools in Coot (v.0.9.4.1)[49], and real_space_refine tool in Phenix (v.1.19.2-4158)[50], using secondary structure, Ramachandran, and rotamer restraints. This complex model served as an initial model for other Cas1–Cas2 structures, which were refined in an analogous manner.

Ligation assays with pCRISPR integration target plasmid

Ligation assays were conducted in integration buffer (20 mM HEPES, pH 7.5, 125 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Nonidet P-40 and 10% DMSO). Cas1 (4 µM) and Cas2/DEDDh (2 µM) were pre-complexed for 30 min at 4 °C before addition of DNA substrate (312.5 nM) and integration target pCRISPR (20 ng ml^{-1}, ~10 nM) and reacting for 2 h at 37 °C. The reaction was quenched with 0.4% SDS and 25 mM EDTA, treated with proteinase K for 15 min at room temperature, and then treated with 3.4% SDS. The reactions were analysed on a 1.5% agarose gel and visualized using the Typhoon FLA gel imaging scanner.

Full-site integration assays

Integration assays (50 µl reactions) were conducted in integration buffer (20 mM HEPES, pH 7.5, 125 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Nonidet P-40 and 10% DMSO). Cas1 (4 µM) and Cas2/DEDDh (2 µM) were pre-complexed for 30 min at 4 °C before addition of DNA substrate containing BsaI cut sites (312.5 nM) and reacting for 15 min, followed by

the addition of the integration target pCRISPR (20 ng ml^{-1}, ~10 nM) and incubating for 2 h at 37 °C. The products were purified using the DNA Clean and Concentrator 5 kit (Zymo Research) and eluted with 6 µl water. A gap-filling reaction (20 µl total, 37 °C for 30 min) was conducted with the purified integration products as described previously[37]: 6 µl purified acquisition reaction, 6.5 µl water, 2 µl 10× Taq DNA ligase buffer (NEB), 2 µl dNTP Solution Mix (10 mM stock, NEB), 2 µl Taq DNA ligase (80 U, NEB) and 1 µl T4 DNA polymerase (1 U, NEB). Gap-filling reactions were purified using the Zymo Research kit and eluted with 6 µl water. A Golden-Gate-compatible chloramphenicol selection cassette was generated by PCR with primers encoding BsaI cut sites and purified using the Qiagen MinElute PCR Purification kit. The sequences of primers used are shown in Supplementary Table 2. A Golden Gate cloning reaction was performed using the purified, gap-filled integration products and chloramphenicol selection cassette according to a standard BsaI assembly protocol. The products were purified using the Zymo Research kit and eluted with 6 µl water, and 1 µl was electroporated into DH10B cells (NEB). Electroporated cells were recovered in 975 µl of LB and plated on LB agar containing carbenicillin (100 µg ml^{-1}) and chloramphenicol (25 µg ml^{-1}). Of the surviving colonies, 95 were sequenced using Sanger sequencing and the sequences were analysed using SnapGene (v.5.0.8).

CRISPR locus bioinformatic analysis

Cas2-DEDDh-containing loci from metagenomic data were identified by determining genomes that contained a CRISPR locus using CRISPRDetect, and coding sequences within 5 kb of the array were extracted[51]. A DEDDh HMM model was built from BLAST searches against the NCBI nr database that were manually verified[52]. The coding sequences were searched against the DEDDh model using hmmsearch with $E < 1 \times 10^{-5}$ (ref. 52). Matches that also contained credible hits to Cas1 and neighbouring other Cas proteins were shortlisted for this work. A preliminary Cas2/DEDDh model was computed using AlphaFold 2 to aid in structure building[47].

Statistics and reproducibility

For biochemical experiments, results represent gels of the highest quality. All experiments were generally performed at least in duplicate, although not in the exact same format. Pilot experiments were performed to ensure reproducibility. Measurements were taken from distinct samples. Full-site integration assays were performed by sequencing 95 colonies and counting integration events in biological triplicate. The choice of sample size was made after ensuring reproducibility through pilot experiments. All data points are displayed on the figure panels.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Atomic models in the Protein Data Bank and the corresponding cryo-EM density maps from the Electron Microscopy Data Bank are available at the following accession codes, respectively: cryo-EM structure of Cas1–Cas2/DEDDh: PAM-deficient prespacer complex (8FY9, EMD-29561); cryo-EM structure of Cas1–Cas2/DEDDh: PAM-containing prespacer complex (8FYA, EMD-29562); cryo-EM structure of Cas1–Cas2/DEDDh: half-site integration complex (8FYB, EMD-29563); cryo-EM structure of Cas1–Cas2/DEDDh: half-site integration complex linear CRISPR repeat conformation (8FYC, EMD-29564); cryo-EM structure of Cas1–Cas2/DEDDh: half-site integration complex with CRISPR repeat bent conformation (8FYD, EMB-29565). The plasmids used in this study are available on reasonable request. Source data are provided with this paper.

References

1. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

2. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*.

Nucleic Acids Res. **40**, 5569–5576 (2012).

3. Nuñez, J. K., Lee, A. S. Y., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature* **519**, 193–198 (2015).

4. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568–571 (2010).

5. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).

6. Hudaiberdiev, S. et al. Phylogenomics of Cas4 family nucleases. *BMC Evol. Biol.* **17**, 232 (2017).

7. Hu, C. et al. Mechanism for Cas4-assisted directional spacer acquisition in CRISPR–Cas. *Nature* **598**, 515–520 (2021).

8. Makarova, K. S. et al. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2019).

9. Kim, S. et al. Selective loading and processing of prespacers for precise CRISPR adaptation. *Nature* **579**, 141–145 (2020).

10. Ramachandran, A., Summerville, L., Learn, B. A., DeBell, L. & Bailey, S. Processing and integration of functionally oriented prespacers in the *Escherichia coli* CRISPR system depends on bacterial host exonucleases. *J. Biol. Chem.* **295**, 3403–3414 (2020).

11. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).

12. Sternberg, S. H., Richter, H., Charpentier, E. & Qimron, U. Adaptation in CRISPR-Cas systems. *Mol. Cell* **61**, 797–808 (2016).

13. McGinn, J. & Marraffini, L. A. Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nat. Rev. Microbiol.* **17**, 7–12 (2019).

14. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663 (2005).

15. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).

16. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).

17. Brouns, S. J. J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).

18. Garneau, J. E. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).

19. Hale, C. R. et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956 (2009).

20. Wright, A. V. & Doudna, J. A. Protecting genome integrity during CRISPR immune adaptation. *Nat. Struct. Mol. Biol.* **23**, 876–883 (2016).

21. Xiao, Y., Ng, S., Nam, K. H. & Ke, A. How type II CRISPR–Cas establish immunity through Cas1–Cas2-mediated spacer integration. *Nature* **550**, 137–141 (2017).

22. Nuñez, J. K., Harrington, L. B., Kranzusch, P. J., Engelman, A. N. & Doudna, J. A. Foreign DNA capture during CRISPR–Cas adaptive immunity. *Nature* **527**, 535–538 (2015).

23. Wang, J. et al. Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. *Cell* **163**, 840–853 (2015).

24. Deveau, H. et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390–1400 (2008).

25. Gleditzsch, D. et al. PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures. *RNA Biol.* **16**, 504–517 (2019).

26. Kieper, S. N. et al. Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation. *Cell Rep.* **22**, 3377–3384 (2018).

27. Lee, H., Zhou, Y., Taylor, D. W. & Sashital, D. G. Cas4-dependent prespacer processing ensures high-fidelity programming of CRISPR arrays. *Mol. Cell* **70**, 48–59 (2018).

28. Drabavicius, G. et al. DnaQ exonuclease-like domain of Cas2 promotes spacer integration in a type I-E CRISPR–Cas system. *EMBO Rep.* **19**, e45543 (2018).

29. Lee, H., Dhingra, Y. & Sashital, D. G. The Cas4-Cas1-Cas2 complex mediates precise prespacer processing during CRISPR adaptation. *eLife* **8**, e44248 (2019).

30. Dhingra, Y., Suresh, S. K., Juneja, P. & Sashital, D. G. PAM binding ensures orientational integration during Cas4-Cas1-Cas2-mediated CRISPR adaptation. *Mol. Cell* **82**, 4353–4367 (2022).

31. Shiriaeva, A. A. et al. Host nucleases generate prespacers for primed adaptation in the *E. coli* type I-E CRISPR–Cas system. *Sci. Adv.* **8**, eabn8650 (2022).

32. Makarova, K. S. & Koonin, E. V. Annotation and classification of CRISPR-Cas systems. *Methods Mol. Biol.* **1311**, 47–75 (2015).

33. Nuñez, J. K. et al. Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **21**, 528–534 (2014).

34. Wright, A. V. et al. Structures of the CRISPR genome integration complex. *Science* **357**, 1113–1118 (2017).

35. Sasnauskas, G. & Siksnys, V. CRISPR adaptation from a structural perspective. *Curr. Opin. Struct. Biol.* **65**, 17–25 (2020).

36. Punjani, A. & Fleet, D. J. 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).

37. Jakhanwal, S. et al. A CRISPR-Cas9–integrase complex generates precise DNA fragments for genome integration. *Nucleic Acids Res.* **49**, 3546–3556 (2021).

38. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).

39. Nuñez, J. K., Bai, L., Harrington, L. B., Hinder, T. L. & Doudna, J. A. CRISPR immunological memory requires a host factor for specificity. *Mol. Cell* **62**, 824–833 (2016).

40. Budhathoki, J. B. et al. Real-time observation of CRISPR spacer acquisition by Cas1-Cas2 integrase. *Nat. Struct. Mol. Biol.* **27**, 489–499 (2020).

41. Shipman, S. L., Nivala, J., Macklis, J. D. & Church, G. M. Molecular recordings by directed CRISPR spacer acquisition. *Science* **353**, aaf1175 (2016).

42. Shipman, S. L., Nivala, J., Macklis, J. D. & Church, G. M. CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature* **547**, 345–349 (2017).

43. Bhattarai-Kline, S. et al. Recording gene expression order in DNA by CRISPR addition of retron barcodes. *Nature* **608**, 217–225 (2022).

44. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

45. Bepler, T. et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).

46. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).

47. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

48. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

49. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

50. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D* **75**, 861–877 (2019).

51. Biswas, A., Staals, R. H. J., Morales, S. E., Fineran, P. C. & Brown, C. M. CRISPRDetect: a flexible algorithm to define CRISPR arrays. *BMC Genom.* **17**, 356 (2016).

52. Steinegger, M. et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinform.* **20**, 473 (2019).

Acknowledgements

We thank the members of the Doudna laboratory for insights; and N. Prywes for reading the manuscript. Financial support was provided by grants from the National Science Foundation. J.Y.W. and G.L. were supported by US National Science Foundation Graduate Research Fellowships (DGE 1752814 and DGE 2146752); and P.S. by the Swiss National Science Foundation Mobility fellowship (P2EZP3_195621). J.A.D. is a Howard Hughes Medical Institute investigator.

Author information

Author notes

1. These authors contributed equally: Joy Y. Wang, Owen T. Tuck, Petr Skopintsev

Authors and Affiliations

1. Department of Chemistry, University of California, Berkeley, CA, USA

   Joy Y. Wang, Owen T. Tuck & Jennifer A. Doudna

2. Innovative Genomics Institute, University of California, Berkeley, CA, USA

   Joy Y. Wang, Owen T. Tuck, Petr Skopintsev, Katarzyna M. Soczek, Gary Li, Basem Al-Shayeb, Julia Zhou & Jennifer A. Doudna

3. Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA

   Petr Skopintsev, Katarzyna M. Soczek, Julia Zhou & Jennifer A. Doudna

4. California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, CA, USA

   Petr Skopintsev, Katarzyna M. Soczek & Jennifer A. Doudna

5. Department of Bioengineering, University of California, Berkeley, CA, USA

   Gary Li & Jennifer A. Doudna

6. Howard Hughes Medical Institute, University of California, Berkeley, CA, USA

Jennifer A. Doudna

7. MBIB Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Jennifer A. Doudna

8. Gladstone Institutes, University of California, San Francisco, CA, USA

Jennifer A. Doudna

Contributions

J.Y.W. conceived the project and designed experiments with O.T.T., P.S., G.L. and J.A.D. J.Y.W., O.T.T., G.L. and J.Z. performed protein purification and in vitro biochemistry. P.S. performed cryo-EM data collection with K.M.S. and O.T.T. P.S., O.T.T. and K.M.S. generated cryo-EM maps and built atomic models with input from J.Y.W. B.A.-S. identified the CRISPR–Cas locus of interest and performed bioinformatics experiments. O.T.T., J.Y.W. and J.A.D. wrote the manuscript with input from all of the authors.

Corresponding author

Correspondence to Jennifer A. Doudna.

Ethics declarations

Competing interests

The Regents of the University of California have patents issued and pending for CRISPR technologies on which J.A.D. is an inventor. J.A.D. is a cofounder of Caribou Biosciences, Editas Medicine, Scribe Therapeutics, Intellia Therapeutics, and Mammoth Biosciences. J.A.D. is a scientific advisory board member of Vertex, Caribou Biosciences, Intellia Therapeutics, Scribe Therapeutics, Mammoth Biosciences, Algen Biotechnologies, Felix Biosciences, The Column Group and Inari. J.A.D. is Chief Science Advisor to Sixth Street, a Director at Johnson & Johnson,

Altos and Tempus, and has research projects sponsored by Apple Tree Partners and Roche.

Peer review

Peer review information

Nature thanks Sandra Garrett and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Cas1—Cas2-DEDDh exhibits strict requirement for pre-duplexed 23 bp DNA duplex prespacer for functional ruler-guided trimming.

(**a**) Prespacer substrates with varying duplex lengths used in processing assay. The duplex and overhang lengths are indicated, and the TT PAM motif is boxed. The arrowheads indicate specific processing positions as observed from the processing products visualized on the denaturing gel in **b**. (**b**) Processing assay with pre-duplexed prespacer substrates shown in **a**. Substrates are schematized. For gel source data, see Supplementary Fig. 1. This experiment was repeated thrice with similar results.

Extended Data Fig. 2 Kinetics of exonuclease trimming and ligation by Cas1:Cas2-DEDDh *in vitro*.

(**a**) Time-course reactions of ruler-guided trimming by Cas1:Cas2-DEDDh using substrates S2 and S3 over two hours. (**b**) Quantification of time-course reactions of ruler-guided trimming shown in **a**. Processing efficiency is

calculated as the percentage of fully processed product formation at 28-29 nt (n = 3 biologically independent experiments). (**c**) Schematic of *in vitro* ligation assay with integration target pCRISPR plasmid and prespacer substrate. Star indicates 6-carboxyfluorescein (6-FAM) label. (**d**) Time-course reactions of ligation assay with substrates from (b) over 2 h. Note formation of a 6-FAM-labelled pCRISPR plasmid ligation product band. For gel source data, see Supplementary Fig. 1. This experiment was repeated twice with similar results.

Source data

Extended Data Fig. 3 Flow-chart of the cryo-EM single particle reconstruction of the PAM-deficient prespacer bound Cas1-Cas2/DEDDh.

(**a**) Map generation pipeline in cryoSPARC consisting of three Topaz training[46] iterations, including representative 2D class averages and 3D maps resulting from *ab initio* reconstruction and further heterogeneous, non-uniform refinement. (**b**) Orientation distribution of the final set of refined particles. (**c**) Gold standard FSC curve of the atomic model refined to the final cryoSPARC sharp map.

Extended Data Fig. 4 Flow-chart of the cryo-EM single particle reconstruction of the PAM-containing prespacer bound Cas1-Cas2 and resolution of the DEDDh density.

(**a**) Cryo-EM data collection parameters and map generation pipeline in cryoSPARC, including representative 2D class averages and 3D maps resulting from *ab initio* reconstruction and further heterogeneous refinement. (**b**) 3D maps, orientation distribution, and gold standard FSC curve of the final cryoSPARC map for the PAM-only (**b**) and DEDDh-containing (**c**) densities. (**d**–**f**) Final sharp maps of Cas1:Cas2-DEDDh bound to prespacers coloured according to local resolution as calculated by cryoSPARC, and the final refined models coloured with B-factors as calculated by Phenix, for: (**d**) PAM-deficient density, (**e**) PAM-containing DEDDh-absent density, and (**f**) PAM-containing, DEDDh-containing density.

Extended Data Fig. 5 Flow-chart of the cryo-EM single particle reconstruction of the PAM-containing half-site substrate bound to Cas1-Cas2/DEDDh.

(**a**) Cryo-EM data collection parameters and map generation pipeline in cryoSPARC, including representative 2D class averages and 3D maps resulting from *ab initio* reconstruction and further heterogeneous refinement. Different particle stacks were used for generation of the high resolution structure (**c**), and 3D variability analysis resulting in linear (**d**) and bent (**e**) density maps. (**b**) Representation of the 4 oligonucleotide half integration construct used in this study. Asterisk indicates a phosphorothioate bond. (**c**–**e**) 3D maps, orientation distribution, and gold standard FSC curve of the atomic model refined to the final cryoSPARC sharp map for each half-site structure.

Extended Data Fig. 6 Incomplete resistance of phosphorothioate PAM nucleotides to DEDDh.

Processing assay (4 h) with substrates shown in above. Substrates are schematized and TT PAM is indicated, along with phosphorothioation. Processing products are depicted on the right. For gel source data, see Supplementary Fig. 1. This experiment was repeated twice with similar results.

Extended Data Fig. 7 Molecular detail of Cas1:Cas2-DEDDh in the linear CRISPR repeat DNA conformation.

(**a**) Orthogonal views of the final sharpened cryo-EM densities for Cas1:Cas2-DEDDh bound to half-site DNA intermediates containing a phosphorothioated TT PAM, coloured to demonstrate domain locations. (**b**) Non-PAM side and bottom views of Cas1:Cas2-DEDDh bound to half-site intermediate DNA. (**c**) Detail depicting potential interaction between Arg^{142} of DEDDh and CRISPR repeat DNA phosphate backbone. (**d**) Non-PAM side view of the DEDDh domain with the sharp map superimposed (threshold: 0.05). Right, detail in the catalytic pocket of DEDDh, with Asp^{132} shown in close proximity to an extension attributed to the N-terminus

of Cas1b. (**e**) Cas1:Cas2-DEDDh linear structure, with protein surfaces except the DEDDh domain coloured according to electrostatic potential.

Extended Data Fig. 8 Integration site preference, orientation bias, and duplex requirements of Cas1—Cas2-DEDDh.

(**a**) Schematic depicting experimental workflow for full-site integration reconstitution. (**b**) Integration events at each integration site from sequenced clones. Integration sites are depicted by arrows at repeat borders. The mean and standard deviation of three independent biological replicates are shown (n = 95 colonies). (**c**) Left, orientation of spacer insertion from prespacer containing PAM motif, + orientation orients original TT PAM containing end toward the leader, − orientation orients original TT PAM containing end away from the leader. Right, Orientation of spacer insertion from control prespacer without PAM motif (the TT PAM from left is replaced with AA). The mean and standard deviation of three independent biological replicates are shown (n = 95 colonies). (**d**) Prespacer substrates with varying duplex lengths used in processing assay. The duplex and overhang lengths are indicated, and the TT PAM motif is boxed. (**e**) Schematic of processing assay with stepwise addition of the top and bottom prespacer strands of the substrates shown in **d**. The unlabelled top PAM-containing strand is incubated first with Cas1:Cas2-DEDDh followed by delayed addition of the labelled bottom PAM-deficient strand. (**f**) Processing assay with stepwise addition of the top and bottom prespacer strands of the substrates shown in **a**. For gel source data, see Supplementary Fig. 1. This experiment was repeated twice with similar results.

Source data

Extended Data Fig. 9 Comparison of Cas1:Cas2-DEDDh bound to the half-site intermediate with the complete *E. coli* IHF-containing integration complex.

(**a**) *E. coli* Cas1:Cas2 DNA IHF holo-complex (PDB: 5WFE). Subunit identities are indicated above. Inset boxes highlight the interface between Cas1a and IHF. (**b**) Representation of the pipeline used to compare *Megasphaera* and *E. coli* integrases. The experimental structure solved in

this work was superimposed with the *E. coli* complex using ChimeraX matchmaker[54]. (**c**) The sequence of a *Megasphaera* IHFα ortholog was found by protein BLAST of *E. coli* IHFα sequence in the *Megasphaera* sp. An286 genome (Taxon ID: 1965622). The AlphaFold 2 sequence of the highest confidence hit is shown along with the protein sequence, with secondary structures indicated. (**d**) The *Megasphaera* IHFα was superimposed with the *E. coli* IHFα to approximate the interface in a hypothetical *Megasphaera* complex. Inset shows the structures overlaid (left) and with only *Megasphaera* proteins shown (right). Right, potential interactions at the interface of *Megasphaera* Cas1 and IHFα.

Extended Data Fig. 10 Comparison of Cas1:Cas2-DEDDh to *E. coli* Cas1:Cas2 structures.

(**a**) *E. coli* Cas1:Cas2 bound to a PAM-containing prespacer (PDB: 5DQZ) was compared to *Megasphaera* Cas1:Cas2-DEDDh bound to a PAM-containing prespacer. The experimental structure solved in this work was superimposed with the *E. coli* complex using ChimeraX matchmaker[54]. (**b**) Inset showing detail of the superimposition at the PAM binding site. The C-terminal loop in *E. coli* Cas1:Cas2 is oriented differently but covers the PAM in an analogous manner. (**a**) *E. coli* Cas1:Cas2 bound to a PAM-absent half-site DNA (PDB: 5VVJ) was compared to *Megasphaera* Cas1:Cas2-DEDDh bound to a PAM-containing half-site DNA. (**d**) Inset depicting differences at the integration strand junction. (**e**) Comparison of the unintegrated side of both holoenzymes.

Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics

Supplementary information

Supplementary Information

Supplementary Fig. 1: lightly annotated raw gel files with boxes indicating cropping for main and extended data figures. Supplementary Tables 1 and 2,

containing the sequences of proteins used in this study, and oligonucleotides used in this study.

Reporting Summary

Supplementary Video 1

Visualization of repeat DNA flexibility of the Cas1:Cas2-DEDDh half-integration complex using 3-D variability analysis. The PAM side is shown. See Materials and Methods for further detail.

Supplementary Video 2

Visualization of repeat DNA flexibility and DEDDh domain dynamics of the Cas1:Cas2-DEDDh half-integration complex using 3-D variability analysis. The bottom view is shown. See Methods for further details.

Source data

Source Data Fig. 4

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 8

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or

exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Wang, J.Y., Tuck, O.T., Skopintsev, P. *et al.* Genome expansion by a CRISPR trimmer-integrase. *Nature* **618**, 855–861 (2023). https://doi.org/10.1038/s41586-023-06178-2

- Received: 26 January 2023
- Accepted: 08 May 2023
- Published: 14 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06178-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06178-2

- Article
- Open Access
- Published: 07 June 2023

Structural basis for FGF hormone signalling

- Lingfeng Chen,
- Lili Fu,
- Jingchuan Sun,
- Zhiqiang Huang,
- Mingzhen Fang,
- Allen Zinkle,
- Xin Liu,
- Junliang Lu,
- Zixiang Pan,
- Yang Wang,
- Guang Liang,
- Xiaokun Li ,
- Gaozhi Chen &
- …
- Moosa Mohammadi

Show authors

Nature **volume 618**, pages 862–870 (2023)

- 6945 Accesses
- 20 Altmetric
- Metrics details

Abstract

α/βKlotho coreceptors simultaneously engage fibroblast growth factor (FGF) hormones (FGF19, FGF21 and FGF23)[1,2] and their cognate cell-surface FGF receptors (FGFR1–4) thereby stabilizing the endocrine FGF–FGFR complex[3,4,5,6]. However, these hormones still require heparan sulfate (HS) proteoglycan as an additional coreceptor to induce FGFR dimerization/activation and hence elicit their essential metabolic activities[6]. To reveal the molecular mechanism underpinning the coreceptor role of HS, we solved cryo-electron microscopy structures of three distinct 1:2:1:1 FGF23–FGFR–αKlotho–HS quaternary complexes featuring the 'c' splice isoforms of FGFR1 (FGFR1c), FGFR3 (FGFR3c) or FGFR4 as the receptor component. These structures, supported by cell-based receptor complementation and heterodimerization experiments, reveal that a single HS chain enables FGF23 and its primary FGFR within a 1:1:1 FGF23–FGFR–αKlotho ternary complex to jointly recruit a lone secondary FGFR molecule leading to asymmetric receptor dimerization and activation. However, αKlotho does not directly participate in recruiting the secondary receptor/dimerization. We also show that the asymmetric mode of receptor dimerization is applicable to paracrine FGFs that signal solely in an HS-dependent fashion. Our structural and biochemical data overturn the current symmetric FGFR dimerization paradigm and provide blueprints for rational discovery of modulators of FGF signalling[2] as therapeutics for human metabolic diseases and cancer.

Main

The mammalian fibroblast growth factor (FGF) family comprises 18 β-trefoil homology-domain-containing polypeptides arranged into five paracrine subfamilies and one endocrine subfamily[7]. Paracrine subfamilies govern multiple events during embryonic development[8] whereas endocrine subfamily members (FGF19, FGF21 and FGF23) are hormones that regulate bile acid, lipid, glucose, vitamin D and mineral ion homeostasis[1,4,9,10]. FGF hormones are promising targets for the treatment of a spectrum of metabolic diseases, including type II diabetes, obesity, non-alcoholic steatohepatitis, primary biliary cirrhosis, bile acid diarrhoea, renal phosphate wasting

disorders and chronic kidney disease[2,11,12,13,14,15,16,17,18,19]. FGFs mediate their actions by binding, dimerizing and thereby activating single-pass transmembrane FGF receptor tyrosine kinases (FGFR1–4)[20,21]. The extracellular region of a prototypical FGFR contains three immunoglobulin (Ig)-like domains (D1, D2 and D3). D2, D3 and the short D2–D3 linker are necessary and sufficient for ligand binding and receptor dimerization. In FGFR1–FGFR3, alternative splicing of two mutually exclusive exons (termed 'b' and 'c') alters the composition of major ligand-binding sites within the D3 domains of these three FGFRs, effectively expanding the number of principal FGFR isoforms to seven (that is, FGFR1b–3b, FGFR1c–3c and FGFR4)[22,23,24].

Paracrine FGFs depend on heparan sulfate (HS) glycosaminoglycans as a mandatory coreceptor to stably bind and dimerize their cognate FGFRs. HS are linear glycan chains of HS proteoglycans (HSPGs) that are abundantly expressed in the extracellular matrix of all tissues[25]. According to the crystal structure of the 2:2:2 FGF2–FGFR1c–HS dimer, HS concurrently engages the HS binding sites of paracrine FGF and FGFR, thereby enforcing FGF–FGFR proximity. In doing so, HS (1) enhances 1:1 FG–FGFR binding affinity; and (2) fortifies interactions between two 1:1 complexes to give rise to two-fold symmetric 2:2 dimers[26]. Extracellular FGFR dimerization promotes the formation of a thermodynamically weak asymmetric complex of the intracellular kinase domains[27] that mediates (A)-loop tyrosine transphosphorylation and hence kinase activation and intracellular signalling.

The HS binding sites of FGF hormones diverge both compositionally and conformationally from those of paracrine FGFs, dramatically weakening their affinity for HS[28]. Consequently, FGF hormones avoid entrapment by the HSPGs in the extracellular matrix and can enter the circulation. Moreover, FGF hormones have weak affinity for FGFRs[28,29] owing to substitutions of their key receptor binding residues. While these structural and biochemical idiosyncrasies confer hormonal mode of action, they render HS insufficient for endocrine FGF to bind FGFR and induce receptor dimerization. Indeed, to offset these deficiencies, FGF hormones have evolved an absolute dependency on αKlotho or βKlotho as an additional coreceptor for signalling. α- and βKlotho are single-pass

transmembrane proteins with a large extracellular domain comprising two tandem glycosidase-like domains (KL1 and KL2), and a short intracellular domain[3,5]. αKlotho (or βKlotho) coreceptor simultaneously binds the FGF hormone and its cognate FGFR, thereby enforcing binding of endocrine FGF to FGFR. Klotho coreceptors have unique FGF hormone and FGFR binding specificities which ultimately dictate FGFR binding specificity and target tissue/organ selectivity of endocrine FGFs. αKlotho binds exclusively to FGF23 (ref. [30]) whereas βKlotho binds both FGF19 and FGF21 (refs. [31,32,33,34]). With respect to FGFR interaction, αKlotho and βKlotho show shared specificity for FGFR1c and FGFR4 but neither recognizes the ‘b’ splice isoforms of FGFR1–3. However, they display opposite specificity towards FGFR2c and FGFR3c with αKlotho binding FGFR3c but not FGFR2c, while βKlotho binds FGFR2c but not FGFR3c (ref. [35]).

The coreceptor mechanism of αKlotho in FGF23 signalling was illuminated by the crystal structure of a ternary complex comprising FGF23, the FGFR1c ligand-binding domain[6] and the soluble ectodomain of αKlotho (a naturally occurring isoform produced via shedding of the transmembrane form)[36,37]. In the structure, the long α1β1 loop (denoted receptor binding arm (RBA)) extends from the KL2 domain of αKlotho and clasps a hydrophobic groove in the D3 domain of FGFR (a conserved hallmark of FGFR1c–3c and FGFR4), while a large cleft at the junction between KL1 and KL2 embraces FGF23’s long C-tail. In doing so, αKlotho enforces FGF23–FGFR proximity and complex stability. As to the role of HS, we previously postulated that HS enables two 1:1:1 FGF23–FGFR–αKlotho complexes to assemble into a symmetric 2:2:2:2 FGF23–FGFR–αKlotho–HS quaternary signalling unit reminiscent of 2:2:2 paracrine FGF–FGFR–HS dimers[6]. To establish the mechanism underlying dual coreceptor (that is, αKlotho and HS) dependent FGFR dimerization/activation by FGF hormones, we solved cryogenic electron microscopy (cryo-EM) structures of all three physiologically possible FGF23–FGFR–αKlotho–HS quaternary complexes. Unexpectedly, these structures, supported by comprehensive cell-based data, reveal that HS fortifies interactions of FGF23 and its primary FGFR within a 1:1:1 FGF23–FGFR–αKlotho with a secondary lone FGFR, thereby inducing asymmetric receptor dimerization and activation. Notably, the asymmetric mode of receptor dimerization is generalizable to

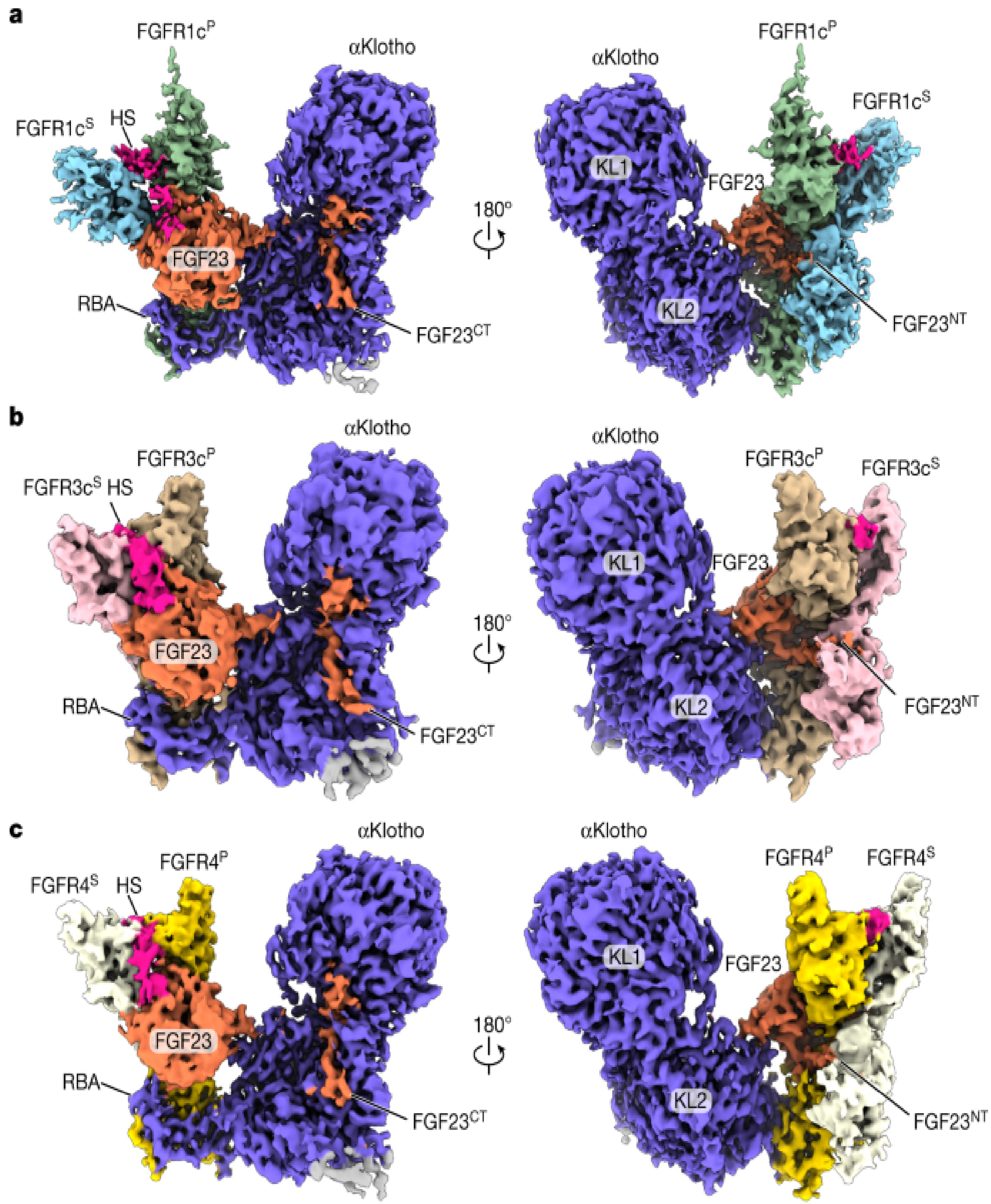
paracrine FGFs, which thus overturns our current symmetric model for FGFR dimerization and activation[26].

1:2:1:1 FGF23–FGFR–αKlotho–HS complexes

We prepared 1:1:1:1 quaternary mixtures of full-length mature human FGF23, the extracellular ligand-binding portions (that is, encompassing D2 and D3 domains) of three human cognate FGFRs of FGF23 (that is, FGFR1c, FGFR3c and FGFR4), whole ectodomain of human αKlotho coreceptor and a fully sulfated heparin dodecasaccharide (hereafter, referred to as HS). The three resulting quaternary complexes (that is, FGF23–FGFR1c–αKlotho–HS, FGF23–FGFR3c–αKlotho–HS and FGF23–FGFR4–αKlotho–HS) were isolated by size-exclusion chromatography (SEC) and were used directly for vitrification, cryo-EM image collection and structure determination (Extended Data Fig. 1, Extended Data Table 1 and Supplementary Fig. 2). Contrary to our prediction, all three cryo-EM structures reveal identical asymmetric 1:2:1:1 FGF23–FGFR–αKlotho–HS quaternary assemblies in which HS enables a 1:1:1 FGF23–FGFR–αKlotho ternary complex to recruit a lone FGFR chain (termed $FGFR^S$ to differentiate it from the ‘primary’ receptor $FGFR^P$ within the ternary complex) (Fig. 1). At the membrane distal end of each quaternary complex, HS engages in tripartite interactions with juxtaposed HS binding sites of FGF23, $FGFR^P$ and $FGFR^S$ (Fig. 2a). In doing so, HS augments interactions of FGF23 and $FGFR^P$ from the FGF23–$FGFR^P$–αKlotho ternary complex with the $FGFR^S$ chain, and thus induces receptor (that is, $FGFR^P$–$FGFR^S$) dimerization (Fig. 3a). Notably, the proximity and orientation (perpendicular to plasma membrane) of C-terminal ends of FGFR chains (approximately 25 Å apart) acquiesces with the formation of an A-loop transphosphorylating asymmetric dimer of intracellular kinase domain[27] (Extended Data Fig. 2a). Although αKlotho does not directly engage $FGFR^S$, it presides over the $FGFR^S$ recruitment by stabilizing the FGF23–$FGFR^P$ complex. Without αKlotho’s assistance, HS cannot independently generate a stable FGF23–$FGFR^P$ complex that is necessary to recruit a secondary $FGFR^S$. Indeed, cell-based experiments confirm that FGF23 signalling is strictly αKlotho dependent (Extended Data Fig. 2b). The conformation of the FGF23–$FGFR^P$–αKlotho ternary complex within the 1:2:1:1 FGF23–FGFR–

αKlotho–HS quaternary assemblies is similar to that of the X-ray structure of HS-free FGF23–FGFR–αKlotho ternary complex (root mean square deviation (RMSD) of only 1.16 Å, Extended Data Fig. 2c)[6]. However, the $FGFR^S$ component of the quaternary complex adopts a distorted conformation that is incompatible with ligand binding, which thus explains the distinct asymmetry of the quaternary complex (Supplementary Fig. 4).

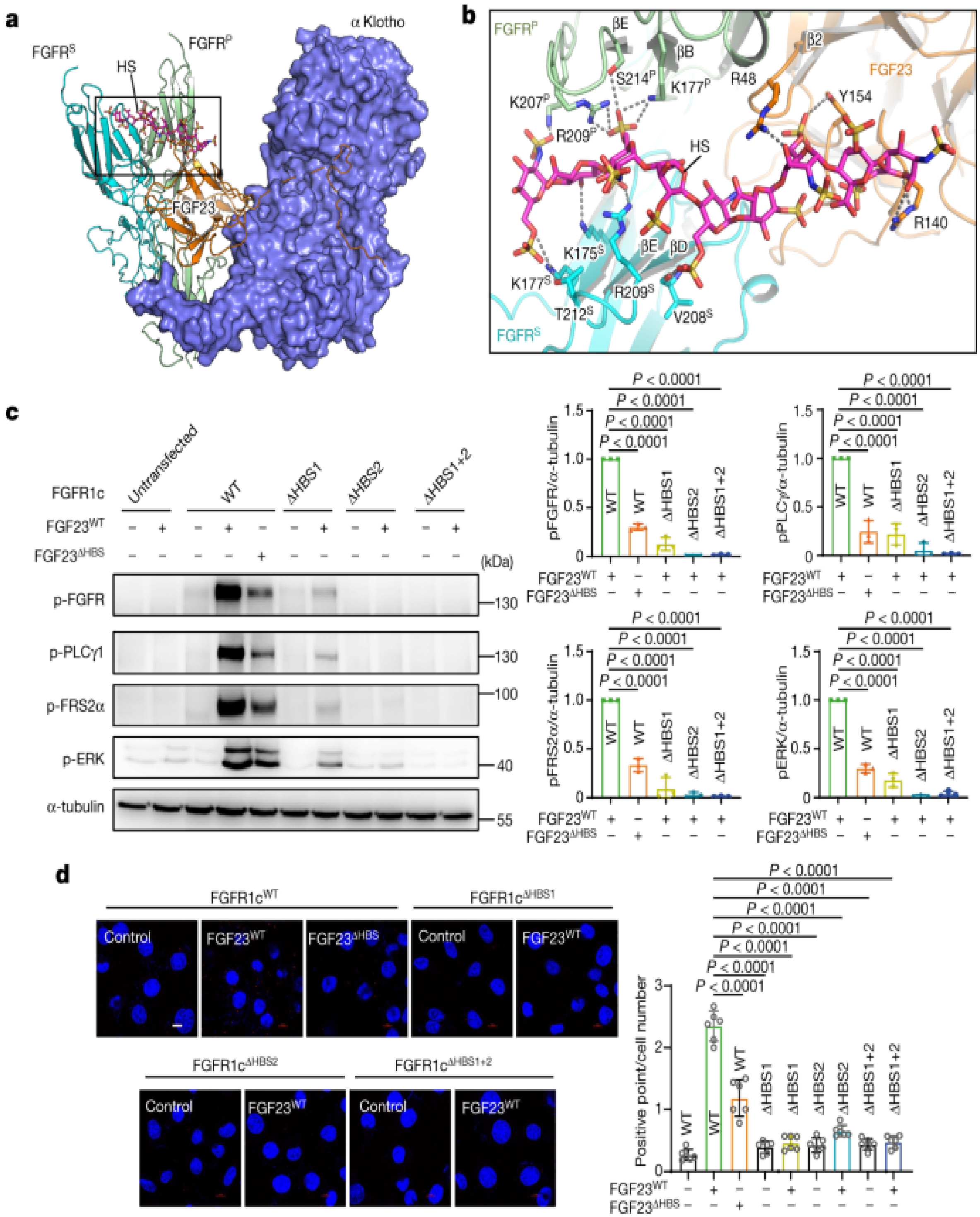
Fig. 1: HS promotes formation of 1:2:1:1 asymmetric FGF23–FGFR–αKlotho–HS quaternary complexes.



Overall view of the cryo-EM reconstructions of FGF23–FGFR1c–αKlotho–HS (**a**), FGF23–FGFR3c–αKlotho–HS (**b**) and FGF23–FGFR4–αKlotho–HS (**c**) quaternary complexes displayed at threshold levels of 0.6, 0.45 and 0.5, respectively. The quaternary complex is shown in two different orientations related by a 180° rotation along the vertical axis. FGF23 is

coloured in orange, αKlotho is shown in deep blue and HS is in magenta. Primary receptor ($FGFR^P$) and secondary receptor ($FGFR^S$) are shown in green and light blue in the FGF23–FGFR1c–αKlotho–HS complex, in tan and pink in the FGF23–FGFR3c–αKlotho–HS complex and in gold and light yellow in the FGF23–FGFR4–αKlotho–HS complex. Two tandem glycosidase-like domains (KL1 and KL2) and RBA of αKlotho are labelled. The relevance of the extra weak density (in grey) seen in all three 1:2:1:1 quaternary complexes is discussed in Supplementary Fig. 3.

Fig. 2: HS promotes asymmetric 1:2:1:1 FGF2–FGFR–αKlotho–HS quaternary complex by simultaneously engaging HBS of FGF23, $FGFR1c^P$ and $FGFR1c^S$.

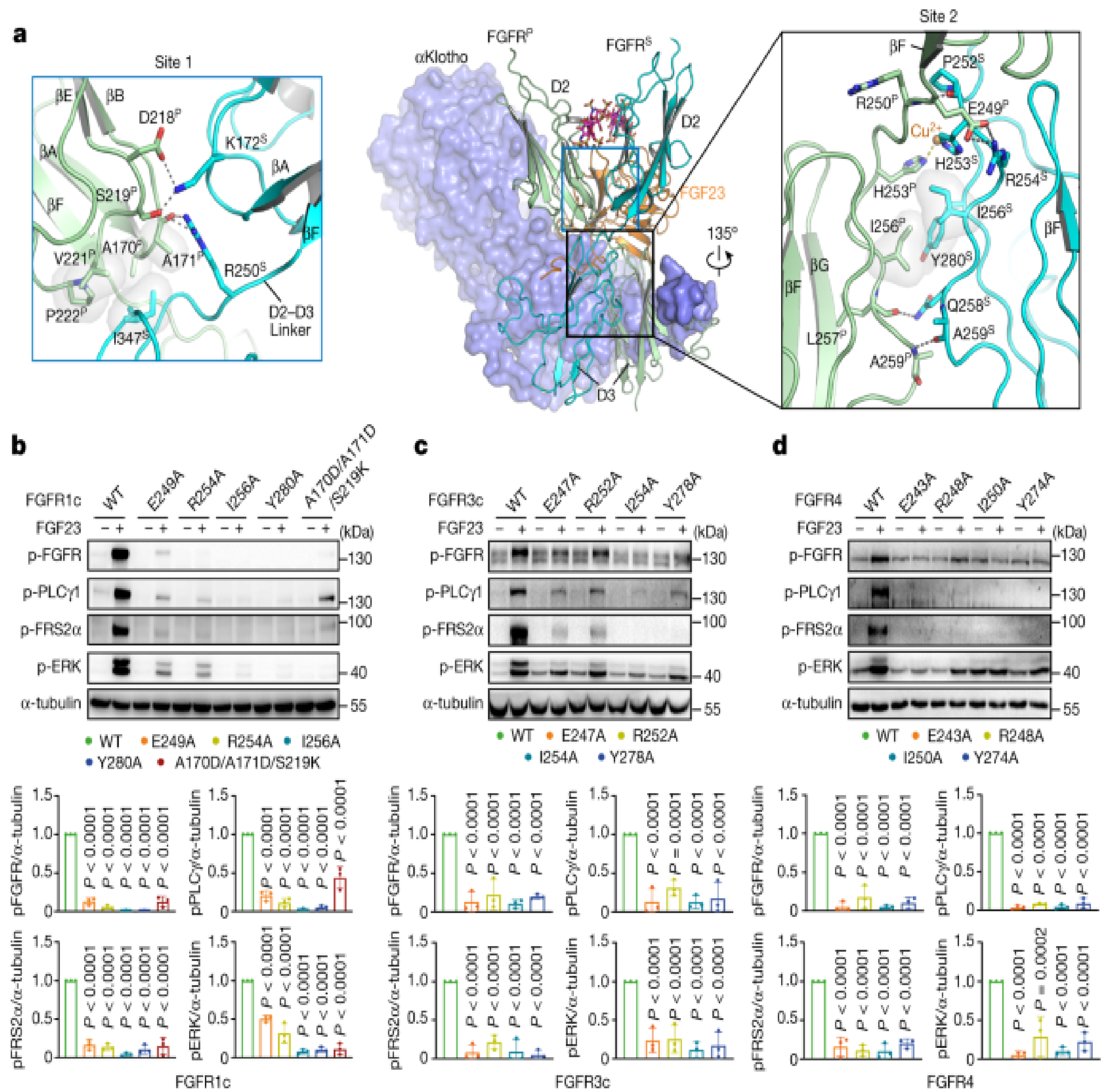


a, FGF23–FGFR1c–αKlotho–HS asymmetric quaternary complex displayed as a hybrid of cartoon (FGF23, FGFR and HS) and surface (αKlotho) in the same orientation as in Fig. 1a (left). **b**, Expanded view of the boxed region in panel **a** showing tripartite interaction of HS with FGF23, $FGFR^P$ and

FGFR^{S}. HS interacting residues are shown as sticks and labelled. Hydrogen bonds in this figure and subsequent figures are represented as black dashed lines. Throughout the figures, nitrogen, oxygen and sulfur atoms are coloured blue, red and yellow, respectively. All structural illustrations were made using Pymol (v.2.5.2). **c**, L6-FGFR1c^{WT}, L6-FGFR1c$^{\Delta HBS1}$, L6-FGFR1c$^{\Delta HBS1}$ or L6-FGFR1c$^{\Delta HBS1+2}$ were cotreated with 20 nM 1:1 mixture of FGF23^{WT} + αKlotho or FGF23$^{R48A/R140A}$ (FGF23$^{\Delta HBS}$) + αKlotho (in the case of L6-FGFR1c^{WT} only) or left untreated. Total cell lysates were immunoblotted with an anti-pY656/Y657–FGFR, anti-pY783-PLCγ1, anti-pY196-FRS2α, anti-pT202/Y204-ERK and anti-β-tubulin antibody (loading control). **d**, Loss in the abilities of FGF23$^{\Delta HBS}$ and FGFR1c$^{\Delta HBS}$ mutants to induce formation of the FGF23–FGFR1c–αKlotho–HS quaternary complex as visualized by PLA. Red spots signify FGF23–FGFR1c–αKlotho–HS quaternary complexes on the cell surface and blue-stained large circles/ovals are cell nuclei. Note that, compared with FGF23^{WT}-treated L6-FGFR1c^{WT} cells, there are far fewer red spots in FGF23$^{\Delta HBS}$-treated L6-FGFRc^{WT} and FGF23^{WT}-treated L6-FGFR1c$^{\Delta HBS}$ cells. Scale bar, 10 μm. Immunoblotting (**c**, normalized against wild-type FGFR1 and FGF23, $n = 3$ biologically independent experiments) and PLA (**d**, $n = 6$ randomly chosen microscope fields from two biologically independent experiments) were quantitated as described in the Methods section and are presented as the mean ± standard deviation. P values were determined by two-way ANOVA followed by Tukey's multiple comparisons post hoc test.

Source data

Fig. 3: Direct FGFR^{P}–FGFR^{S} contacts are required for receptor dimerization and activation.



a, Representation of the FGF23–FGFR1c–αKlotho–HS complex structure as a mix of cartoon (FGF23 and FGFR) and surface (αKlotho). View is related to that in Fig. 1a (right) by 90° rotation along the vertical axis. Contact sites 1 and 2 of the $FGFR1c^{P}$–$FGFR1c^{S}$ dimer interface are boxed in blue and black, respectively. Note that αKlotho does not directly participate in recruiting $FGFR1^{S}$. Left, magnified view of site 1 involving D2 domain of $FGFR1c^{P}$ and D2, D2–D3 linker and D3 of $FGFR1c^{S}$. Right, close-up view of site 2 between D3 domains of $FGFR1c^{P}$ and $FGFR1c^{S}$. Side chains of the interacting residues are shown as sticks. Selected secondary structure elements are labelled. Hydrophobic contacts are highlighted as a semitransparent surface. A Cu^{2+} ion (orange sphere) is coordinated by

analogous histidine residues from FGFR1c^P and FGFR1c^S. Cu^{2+} assignment was based on a previous publication implicating specific Cu^{2+} interactions with extracellular domains of FGFRs[39]. Cu^{2+} ions probably derive from cell culture media (DMEM and DME/F12) used to grow HEK293S GnTI$^-$ cells that secrete minimally glycosylated FGFR ectodomains. **b**–**d**, Immunoblot analyses of whole cell extracts probed as in Fig. 2c from unstimulated or FGF23 (20 nM) and αKlotho (20 nM) costimulated FGFR1c (**b**), FGFR3c (**c**) or FGFR4 (**d**) expressed L6 cell lines. Immunoblotting data were quantitated as described in the Methods section and are presented as the mean ± standard deviation. Data normalized against wild-type FGFR and FGF23, $n = 3$ biologically independent experiments. P values were determined by two-way ANOVA followed by Tukey's multiple comparisons post hoc test.

Source data

Tripartite interactions of HS

FGF23–HS interactions are mediated by three residues (Arg48, Arg140 and Tyr154) from the atypical HS binding site within the core of FGF23 (Fig. 2b). Compared with FGF23–HS contacts, HS interacts more extensively with HBS in the D2 domains of the two FGFR chains, involving residues Lys177, Lys207, Arg209 and Ser214 of FGFR1c^P and Lys175, Lys177, Val208, Arg209 and Thr212 of FGFR1c^S (Fig. 2b). To validate the physiological importance of three-way interactions of HS with FGF23, FGFR1c^P and FGFR1c^S for the induction of receptor (that is, FGFR^P–FGFR^S) dimerization, we introduced an R48A/R140A double mutation into FGF23 (FGF23$^{\Delta HBS}$) and K175Q/K177Q and K207Q/R209Q double mutations (FGFR1c$^{\Delta HBS1}$ and FGFR1c$^{\Delta HBS2}$) separately or in combination (K175Q/K177Q/K207Q/R209Q termed FGFR1c$^{\Delta HBS1+2}$) into the full-length FGFR1c. Wild-type FGFR1c (FGFR1c^{WT}) and its mutated variants were stably expressed on the surface of rat skeletal myoblast cells (L6), an αKlotho- and FGFR-deficient cell line. Endoglycosidase H (Endo H) and Peptide:N-glycosidase F (PNGase F) treatment of cell lysates followed by immunoblot analysis with FGFR-specific antibodies showed that mutated FGFR1c variants contain complex sugars, which implies that they reside on

the cell surface (Extended Data Fig. 3a). Equal homogeneity/quantity of $FGF23^{WT}$ and $FGF23^{\Delta HBS}$ samples was verified by SDS–PAGE (Supplementary Fig. 2d). Cotreatment with $FGF23^{WT}$ and soluble αKlotho of L6-$FGFR1c^{WT}$ led to robust FGFR1c activation/signalling as measured by phosphorylation of FGFR1c on A-loop tyrosines, its two direct substrates PLCγ1 (on the regulatory Y783) and FRS2α (on the Grb2 recruitment site Y196) and subsequent activation of a RAS–mitogen-activated protein kinase (MAPK) pathway as monitored by extracellular signal-regulated kinase (ERK) phosphorylation on T202/Y204. In contrast, both $FGFR1c^{\Delta HBS1}$ and $FGFR1c^{\Delta HBS2}$ double mutants suffered major losses in their ability to induce FGF23 signalling, and the $FGFR1c^{\Delta HBS1+2}$ quadruple mutant became totally silent (Fig. 2c). Likewise, $FGF23^{\Delta HBS}$ was significantly retarded in its ability to activate L6-$FGFR1c^{WT}$ in the presence of soluble αKlotho. FGFR activation/signalling data were mirrored by proximity ligation assay (PLA) data. Specifically, cotreatment with $FGF23^{WT}$ and αKlotho led to appearance of copious and intense punctate fluorescent signals on the surface of the L6-$FGFR1c^{WT}$ cell line. In contrast, there were far fewer fluorescent signals on the surface of $FGFR1c^{\Delta HBS1}$, $FGFR1c^{\Delta HBS2}$ and $FGFR1c^{\Delta HBS1+2}$ cell lines upon cotreatment. Similarly, markedly fewer fluorescent dots were present on the surface of the L6-$FGFR1c^{WT}$ cell line when cotreated with $FGF23^{\Delta HBS}$ and soluble αKlotho (Fig. 2d). These cell-based data confirm the indispensability of FGF23–HS and FGFR1c–HS contacts in inducing formation of a FGF23–FGFR1c–αKlotho–HS quaternary cell-surface signalling complex. Consistent with the 1:2:1:1 FGF23–FGFR–αKlotho–HS stoichiometry in the cryo-EM structures, size-exclusion chromatography–multi-angle light scattering experiments showed that the FGF23–FGFR1c–αKlotho–HS quaternary complex migrates as a single species with a calculated molecular mass of approximately 220 kDa, which matches closely the theoretical value for a 1:2:1:1 FGF23–FGFR1c–αKlotho–HS (215 kDa) (Supplementary Fig. 2e).

Direct $FGFR^P$–$FGFR^S$ interactions

All three subdomains (that is, D2, D2–D3 linker and D3) of both $FGFR^P$ and $FGFR^S$ chains take part in recruitment of $FGFR^S$ to the ternary complex and

hence in promoting receptor dimerization (Fig. 3a). The direct $FGFR^P$–$FGFR^S$ contacts bury a modest solvent-exposed surface area (1542.6 Å^2) and are preserved among the three cryo-EM structures. The $FGFR^P$–$FGFR^S$ interface can be split into two sites (that is, sites 1 and 2). At site 1, residues from discontinuous loop regions (that is, βA'–βB and βE–βF loops) at the bottom (C-terminal) corner of D2 of $FGFR^P$ asymmetrically pack against D2, D2–D3 linker and D3 of the $FGFR^S$ (Fig. 3a). At site 2, a contiguous stretch of residues encompassing D2–D3 linker, βA strand, βA–βA' loop and βB' strand of D3, engage the corresponding region of the secondary receptor in a pseudo-symmetric fashion (Fig. 3a and Supplementary Table 1). Notably, the $FGFR^P$–$FGFR^S$ interface traps two crucial ligand-binding residues of $FGFR^S$, namely, Arg250 and Ser346, which further deprives $FGFR^S$ from ligand binding (Supplementary Fig. 4b).

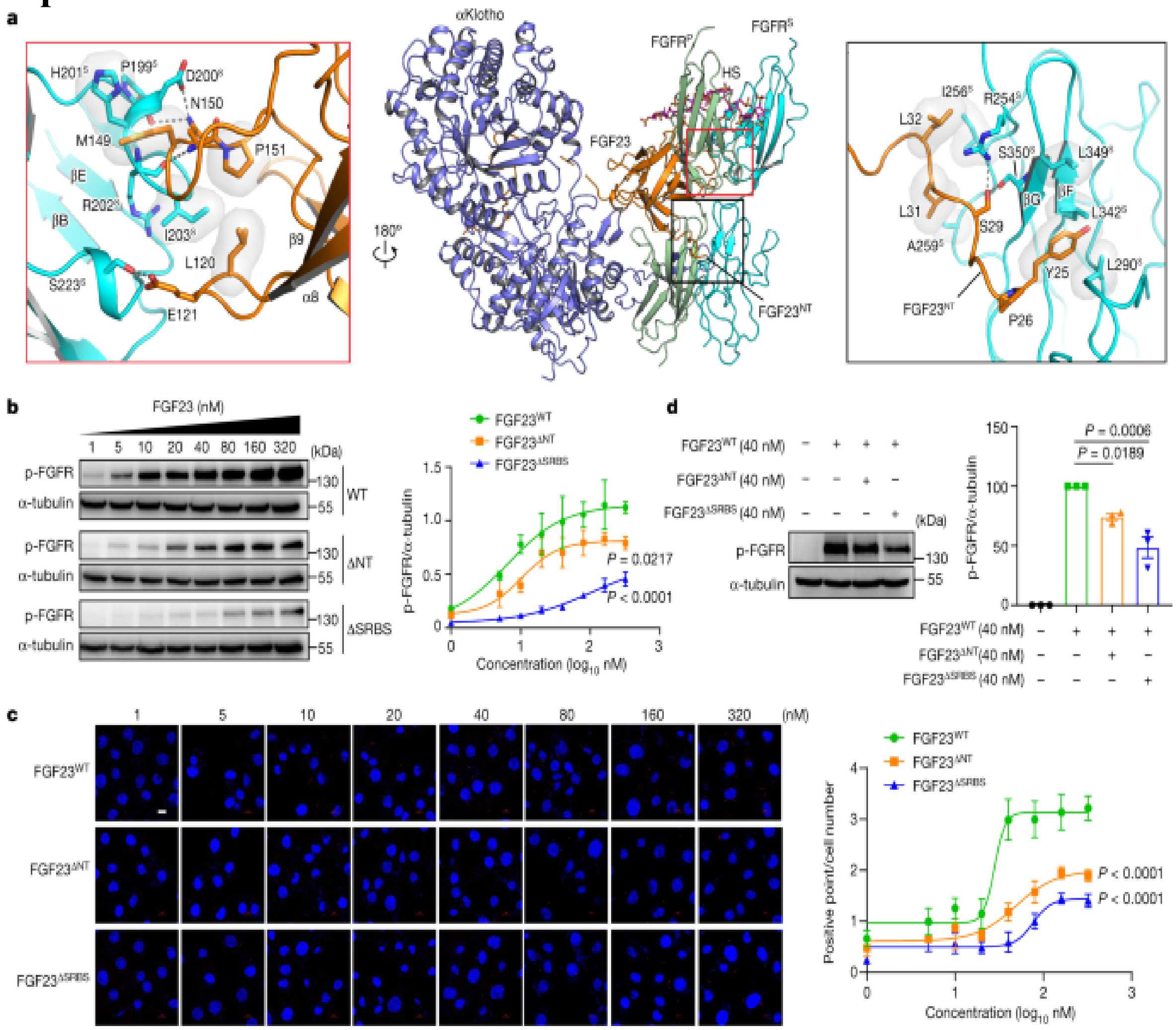
We targeted site 2 of the $FGFR^P$–$FGFR^S$ interface in each of the three quaternary complexes by introducing four different point mutations individually into full-length FGFR1c (E249A, R254A, I256A, Y280A), FGFR3c (E247A, R252A, I254A, Y278A) or FGFR4 (E243A, R248A, I250A, Y274A). In addition, we introduced a A170D/A171D/S219K triple mutation into FGFR1c to target site 1 of the $FGFR1c^P$–$FGFR1c^S$ dimer interface (Fig. 3a) as the representative of the three quaternary complexes. Full-length FGFR1c, FGFR3c and FGFR4 variants along with their wild-type counterparts were stably expressed in L6 cells. As above, Endo H sensitivity was used to ensure that the mutations do not affect receptor glycosylation/maturation and cell-surface presentation (Extended Data Fig. 3b–d). Apart from the $FGFR1c^{I256A}$ and its corresponding $FGFR4^{I250A}$ mutant, all mutants showed comparable ratios of Endo H-resistant to total FGFR protein-like wild-type FGFRs. These data imply that only $FGFR1c^{I256A}$ and corresponding $FGFR4^{I250A}$ had reduced cell-surface expression. As in the case of the L6-$FGFR1c^{WT}$ cell line, cotreatment with FGF23 and soluble αKlotho of L6-$FGFR3c^{WT}$ and L6-$FGFR4^{WT}$ cell lines robustly stimulated FGFR signalling as detected by phosphorylation/activation of FGFRs and downstream PLCγ1/FRS2α/ERK signal transducers. In contrast, cells expressing mutated FGFR1c, FGFR3c and FGFR4 had diminished FGFR A-loop tyrosine phosphorylation and reduced PLCγ1, FRS2α and MAPK activation (Fig. 3b–d). It is unlikely that

the impaired signalling by $FGFR1c^{I256A}$ and $FGFR4^{I250A}$ is solely due to reduced cell-surface expression because the glycosylation and cell-surface expression of the corresponding $FGFR3c^{I254A}$ is unaffected.

Secondary FGF23–$FGFR^S$ contacts

FGF23 also participates in recruiting $FGFR^S$ to the ternary complex and hence in buttressing $FGFR^P$–$FGFR^S$ dimerization. These secondary FGF23–$FGFR^S$ contacts are mediated by both the rigid trefoil core and the flexible N terminus of FGF23. Specifically, the β8–β9 and β10–β12 loops within FGF23's trefoil core engage the βC–βD and βE–βF loops at the bottom edge (C-terminal) of the $FGFR^S$ D2 domain (Fig. 4a). In parallel, hydrophobic residues at the very distal end of the FGF23 N terminus extend from the FGF23–$FGFR^P$–αKlotho complex and interact with a hydrophobic groove in the $FGFR^S$ D3 domain that corresponds to the D3 groove in $FGFR^P$ engaged by the RBA of αKlotho (Fig. 4a and Extended Data Fig. 4a). In doing so, the FGF23 N terminus is likely to discourage binding of another αKlotho to $FGFR^S$ and hence contributes to the quaternary complex asymmetry. When compared with the crystal structure of the HS-free 1:1:1 FGF23–FGFR–αKlotho complex, the N terminus of FGF23 displays a major conformational change in the 1:2:1:1 FGF23–FGFR–αKlotho–HS quaternary complex (Extended Data Fig. 4b–d). Notably, in all three cryo-EM structures, electron densities for the FGF23 N terminus are poorly defined, which implies that the FGF23 N terminus engages $FGFR^S$ D3 in a rather degenerate and flexible fashion (Fig. 1a–c and Extended Data Fig. 4e). Interactions of the FGF23 N terminus with $FGFR^S$ D3 were corroborated by a 300 ns all-atom molecular dynamics (MD) simulation of the FGF23–FGFR1c–αKlotho–HS complex (Extended Data Fig. 4f–i). Akin to the $FGFR^P$–$FGFR^S$ interface, the FGF23–$FGFR^S$ interface (Supplementary Table 2) is also conserved among the three quaternary complexes and masks a rather modest surface exposed area (1668.6 $Å^2$). Notably, the FGF23 N terminus and D3 of the $FGFR^S$ engage each other in the proximity of site 2 of the $FGFR^P$–$FGFR^S$ interface (Fig. 4a). This implies that FGF23–$FGFR^S$ and $FGFR^P$–$FGFR^S$ contacts act in concert to recruit $FGFR^S$ to the ternary complex and promote receptor dimerization (that is, $FGFR^P$–$FGFR^S$).

Fig. 4: Secondary contacts between FGF23 and FGFR^S are essential for receptor dimerization and activation.



a, Cartoon representation of the FGF23–FGFR–αKlotho–HS structure in the same view as in Fig. 1a (right). Orange and black boxes signify the two contact regions between FGF23 and $FGFR^S$, namely, $FGF23^{core}$:$FGFR1c^S$ D2 domain (left, orange box) and $FGF23^{NT}$:$FGFR1c^S$ D3 domain (right). **b**, L6-$FGFR1c^{WT}$ cells were treated with increasing concentrations of $FGF23^{WT}$, $FGF23^{\Delta NT}$ or $FGF23^{\Delta SRBS}$ and whole cell lysates were probed as in Fig. 2. Equal homogeneity/quantity of $FGF23^{WT}$, $FGF23^{\Delta NT}$ and $FGF23^{\Delta SRBS}$ samples was verified by SDS–PAGE (Supplementary Fig. 2d). **c**, Loss in the abilities of $FGF23^{\Delta NT}$ and $FGF23^{\Delta SRBS}$ mutants to induce FGF23–FGFR–αKlotho–HS quaternary complex formation (that is, receptor dimerization) relative to $FGF23^{WT}$ as assessed by PLA. Note that there are

far fewer red spots in cells stimulated with FGF23 mutants relative to $FGF23^{WT}$. **d**, L6-$FGFR1c^{WT}$ cells were treated with $FGF23^{WT}$ (40 nM) alone, $FGF23^{WT}$ (40 nM) mixed with $FGF23^{\Delta SRBS}$ (40 nM) or $FGF23^{\Delta NT}$ (40 nM) and whole cell extracts were probed with pFGFR antibodies. Scale bar, 10 μm. Immunoblotting (**b** and **d**, $n = 3$ biologically independent experiments) and PLA (**c**, $n = 6$ randomly chosen microscope fields from two biologically independent experiments) were quantitated as described in the Methods section and are presented as the mean ± standard deviation. P values were determined by two-way ANOVA followed by Tukey's multiple comparisons post hoc test.

Source data

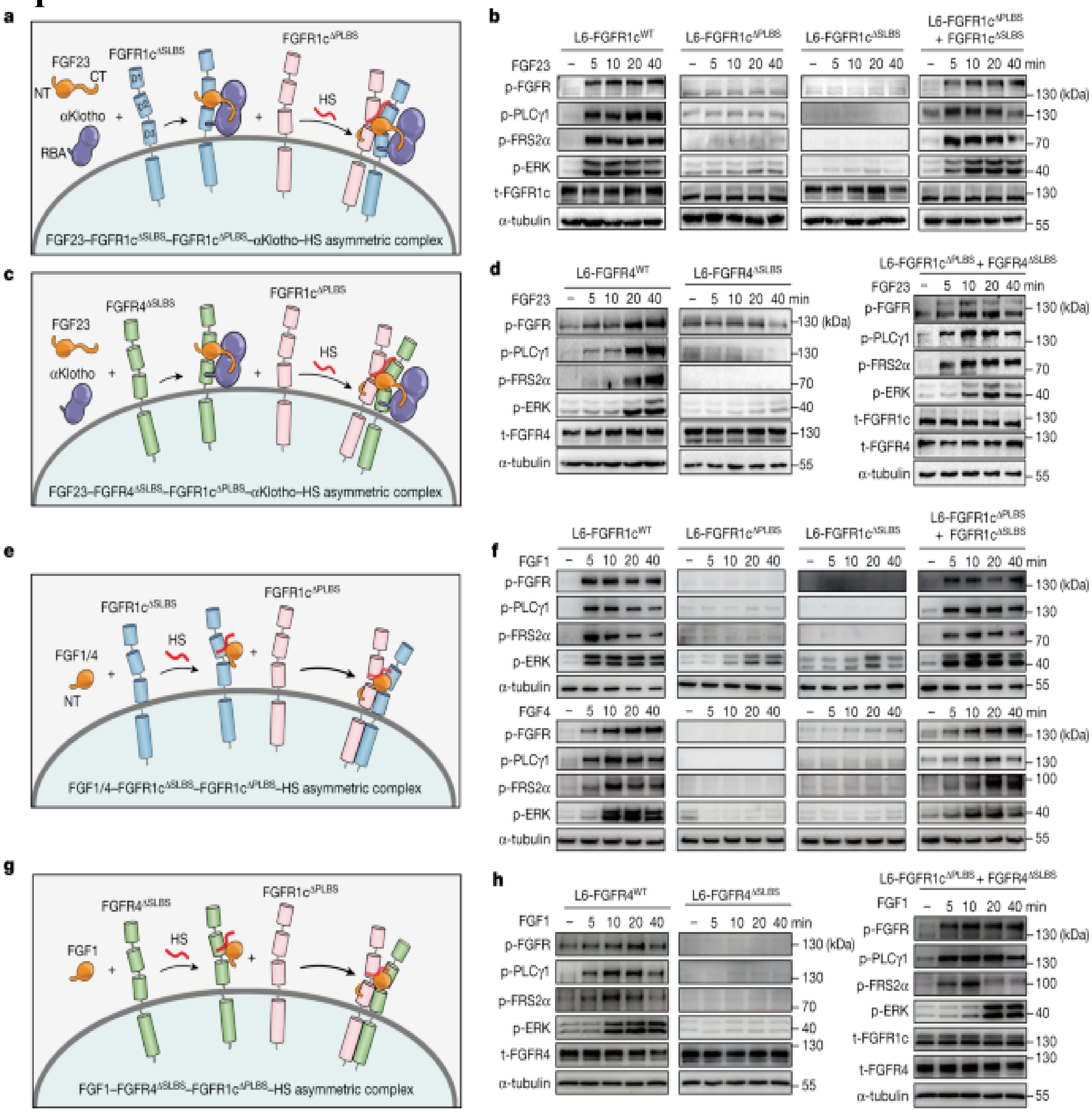
Next, we studied the impact of interfering with the secondary FGF23–$FGFR^S$ contacts on formation of the 1:2:1:1 FGF23–FGFR–αKlotho–HS quaternary signalling complex. For this purpose, we generated two FGF23 variants: one lacking its first 12 N-terminal residues (that is, Y25 to W36; termed $FGF23^{\Delta NT}$), and the other harbouring a M149A/N150A/P151A triple mutation within the core region of the FGF23 (termed $FGF23^{\Delta SRBS}$). According to the cryo-EM structures, the M149A/N150A/P151A triple mutation and N-terminal truncation are predicted to abrogate secondary interactions of FGF23 with D2 and D3 domains of $FGFR^S$, respectively, and hence impair FGF23 signalling (Fig. 4a). To test this, an L6 cell line co-expressing $FGFR1c^{WT}$ and transmembrane αKlotho was generated and exposed to increasing concentrations of $FGF23^{WT}$, $FGF23^{\Delta NT}$ or $FGF23^{\Delta SRBS}$. Receptor dimerization/activation efficacies of increasing concentrations of ligands were assessed by immunoblotting for FGFR A-loop tyrosine phosphorylation and by PLA (Fig. 4b,c). Comparison of dose–response curves showed that, at all concentrations, neither $FGF23^{\Delta NT}$ nor $FGF23^{\Delta SRBS}$ could reach the maximal activity (E_{max}) exerted by $FGF23^{WT}$ in both assays. Moreover, when combined with $FGF23^{WT}$, both $FGF23^{\Delta SRBS}$ and $FGF23^{\Delta NT}$ acted as competitive antagonists producing a net decrease in FGFR1c activation (Fig. 4d). However, relative to $FGF23^{\Delta NT}$, $FGF23^{\Delta SRBS}$ showed a greater loss in its dimerization/signalling capacity (Fig. 4b–d), which implies that the FGF23–$FGFR^S$ contacts mediated via the core of the FGF23 contribute more to the overall stability and functionality of the

FGF23–FGFR1c–αKlotho–HS quaternary signalling complex than those mediated by FGF23 N terminus.

Cell-based receptor complementation

We devised a cell-based receptor complementation assay to interrogate the asymmetry of FGF23–FGFR–αKlotho–HS quaternary signalling complexes. Specifically, based on the distinct roles played by primary and secondary receptors in the asymmetric dimerization, we engineered two FGFR1c variants capable of acting exclusively either as the primary or secondary receptor. We reasoned that although these variants would be dysfunctional individually, they should complement each other and form a functional 1:2:1:1 quaternary complex in response to FGF 23 and αKlotho (Fig. 5a and Extended Data Fig. 5a,b). To generate a FGFR1c variant functioning exclusively as primary receptor, we introduced a I203E/S223E double mutation in its D2 domain. As mentioned above, these two residues in $FGFR1c^{S}$ engage in hydrophobic and hydrogen bonding contacts with residues in the core of the FGF23 (that is, FGF23–$FGFR1c^{S}$ interface, Fig. 4a), which are indispensable for recruitment of $FGFR1c^{S}$ to the ternary complex as implied by the severe loss of function of $FGF23^{\Delta SRBS}$ (Fig. 4b,c). However, the corresponding residues in $FGFR1c^{P}$ are dispensable for quaternary complex formation and are solvent exposed (Extended Data Fig. 5b). Consequently, the resulting mutant FGFR1c (termed $FGFR1c^{\Delta SLBS}$) is predicted to lose the ability to act as a secondary receptor but should retain the ability to serve as a primary receptor. To make an FGFR1c variant capable of acting solely as secondary receptor, we introduced a A167D/V248D double mutation into its D2 domain. The A167D/ V248D double mutation abrogates formation of highly conserved hydrogen and hydrophobic contacts between FGF23 and the $FGFR1^{P}$ D2 domain (Extended Data Fig. 5a,b). Thus, the resulting $FGFR1c^{A167D/V248D}$ mutant (termed $FGFR1c^{\Delta PLBS}$) is predicted to lose the ability to function as a primary receptor (that is, form a ternary complex). However, $FGFR1c^{\Delta PLBS}$ should retain its ability to function as a secondary receptor because the corresponding residues in $FGFR1c^{S}$ do not play any role in the quaternary complex formation and are completely solvent exposed.

Fig. 5: Both endocrine and paracrine FGFs signal via asymmetric receptor dimers.



a, Schematic diagram showing that, in response to FGF23 and αKlotho cotreatment, FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ can complement each other and form a 1:1:1:1:1 FGF23–FGFR1c$^{\Delta SLBS}$–FGFR1c$^{\Delta PLB}$–αKlotho–HS asymmetric complex. **b**, Immunoblot analyses of whole extracts from untreated or FGF23- and αKlotho-cotreated L6 cell lines singly expressing FGFR1c^{WT}, FGFR1c$^{\Delta PLBS}$ and FGFR1c$^{\Delta SLBS}$ or co-expressing FGFR1c$^{\Delta SLBS}$ with FGFR1c$^{\Delta PLBS}$ probed as in Fig. 2. **c**, Schematic diagram showing that FGF23, αKlotho and FGFR4$^{\Delta SLBS}$ (serving as a primary

receptor) form a ternary complex and recruit FGFR1c$^{\Delta PLBS}$ as a secondary receptor, in the presence of HS. **d**, FGF23- and αKlotho-cotreated L6 cell lines stably expressing FGFR4^{WT}, FGFR4$^{\Delta SLBS}$ alone or FGFR4$^{\Delta SLBS}$ together with FGFR1c$^{\Delta PLBS}$ were analysed for FGFR activation/signalling using western blotting of total cell lysates as in Fig. 2. **e**, Schematic diagram showing that, in response to paracrine FGF1/4 and HS, FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ can complement each other and form 1:1:1:1 FGF–FGFR1c$^{\Delta SLBS}$–FGFR1c$^{\Delta PLBS}$–HS asymmetric signalling complexes. **f**, L6 cell lines expressing FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ individually or co-expressing them were treated with FGF1 (1 nM) and cell extracts were immunoblotted as in Fig. 2. **g**, Schematic diagram showing that FGF1, HS and FGFR4$^{\Delta SLBS}$ (serving as primary receptor) form a stable complex which subsequently recruits FGFR1c$^{\Delta PLBS}$ as a secondary receptor. **h**, L6 cell lines expressing FGFR4^{WT} or FGFR4$^{\Delta SLBS}$ singly or co-expressing FGFR4$^{\Delta SLBS}$ with FGFR1c$^{\Delta PLBS}$ were treated with FGF1 (1 nM) and cell extracts were immunoblotted as in Fig. 2. Experiments were performed in biological triplicates with similar results. CT, C terminus; NT, N terminus.

To perform the complementation assay, an L6 cell line co-expressing FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ (L6-FGFR1c$^{\Delta SLBS}$ + FGFR1c$^{\Delta PLBS}$) was generated along with two control L6 cell lines individually expressing FGFR1c$^{\Delta SLBS}$ (L6-FGFR1c$^{\Delta SLBS}$) and FGFR1c$^{\Delta PLBS}$ (L6-FGFR1c$^{\Delta PLBS}$). These three cell lines and L6-FGFR1c^{WT} (positive control) were exposed to a fixed concentration of FGF23 and soluble αKlotho for increasing time intervals. FGFR1c signalling was assessed by monitoring tyrosine phosphorylation of FGFR, PLCγ1 and FRS2α, and subsequent activation of MAPK pathway. Compared with the L6-FGFR1c^{WT} cell line, there was negligible FGFR signalling in L6-FGFR1c$^{\Delta PLBS}$ or L6-FGFR1c$^{\Delta SLBS}$ cells (Fig. 5b), which implies that these variants on their own are unable to form signalling-competent quaternary complexes. In contrast, co-expression of FGFR1c$^{\Delta SLBS}$ with FGFR1c$^{\Delta PLBS}$ (Fig. 5b) resulted in robust activation of an FGFR signalling pathway. These data imply that FGFR1c$^{\Delta SLBS}$ can complement FGFR1c$^{\Delta PLBS}$ in forming a signalling-competent FGF23–FGFR1c$^{\Delta SLBS}$–FGFR1c$^{\Delta PLBS}$–αKlotho–HS quaternary complex, in which FGFR1c$^{\Delta SLBS}$ assumes the role of primary receptor while FGFR1c$^{\Delta PLBS}$ is adopted as the secondary receptor (Fig. 5a). Formation of an asymmetric

FGF23–FGFR1c$^{\Delta SLBS}$–FGFR1c$^{\Delta PLBS}$–αKlotho–HS signalling complex was also validated by PLA (Extended Data Fig. 5c). Compared with L6-FGFR1c^{WT} cells, FGF23 and αKlotho cotreatment failed to generate any fluorescent signals on the surface of L6-FGFR1c$^{\Delta SLBS}$ and L6-FGFR1c$^{\Delta PLBS}$ cells, which implies that FGFR1c$^{\Delta SLBS}$ and L6-FGFR1c$^{\Delta PLBS}$ variants alone fail to undergo receptor dimerization (that is, quaternary complex formation). In contrast, stimulation of L6-FGFR1c$^{\Delta SLBS}$ + FGFR1c$^{\Delta PLBS}$ co-expressing cells with FGF23 and αKlotho led to appearance of copious and intense florescent puncta, implying that FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ have complemented each other and assembled into an asymmetric FGF23–FGFR1c$^{\Delta SLBS}$–FGFR1c$^{\Delta PLBS}$–αKlotho–HS signalling complex.

Cell-based receptor heterodimerization

As mentioned above, both FGFR^{P}–FGFR^{S} and FGF23–FGFR^{S} interfaces are nearly invariant among the three quaternary complexes. This observation afforded us with another opportunity to interrogate the asymmetry of the complex. Specifically, we reasoned that a primary FGFR1c^{P} should be able to pair with FGFR3c or FGFR4 as the secondary receptor to yield functional 1:2:1:1 FGF23–FGFR1c–FGFR3c–αKlotho–HS or FGF23–FGFR1c–FGFR4–αKloth–HS heterodimeric quaternary complexes (Fig. 5c). To test this possibility, we carried out a receptor complementation assay using an FGFR4 variant harbouring a double I197E/S217E mutation (termed FGFR4$^{\Delta SLBS}$)—corresponding to FGFR1c$^{\Delta SLBS}$—as the primary receptor and FGFR1c$^{\Delta PLBS}$ as the secondary receptor. An L6 cell line co-expressing FGFR1c$^{\Delta PLBS}$ with FGFR4$^{\Delta SLBS}$ (L6-FGFR1c$^{\Delta PLBS}$ + FGFR4$^{\Delta SLBS}$), along with control cell lines individually expressing FGFR4^{WT} (L6-FGFR4^{WT}) and FGFR4$^{\Delta SLBS}$ (L6-FGFR4$^{\Delta SLBS}$) were derived. These cell lines were cotreated with FGF23 plus αKlotho, and quaternary signalling complex formation was examined by immunoblot analysis and PLA. As with FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$, FGFR4$^{\Delta SLBS}$ also failed to be activated in response to FGF23 plus αKlotho cotreatment. However, robust activation of an FGFR signalling pathway took place in the L6-FGFR1c$^{\Delta PLBS}$ + FGFR4$^{\Delta SLBS}$ co-expressing cell line (Fig. 5d). The cell signalling data were

mirrored by the PLA (Extended Data Fig. 5c). Specifically, inconsequential florescent signal was present on the surface of L6-FGFR4$^{\Delta SLBS}$ cell line upon costimulation with FGF23 and αKlotho. In contrast, copious and intense punctate fluorescent spots appeared on the surface of L6-FGFR1c$^{\Delta PLBS}$ + FGFR4$^{\Delta SLBS}$ co-expressing cells, which implies that an FGF23–FGFR1c–FGFR4–αKlotho–HS heterodimeric quaternary complex can form on the surface of live cells. These cell-based receptor complementation and hererodimerization data unequivocally confirm the asymmetry of our structurally deduced 1:2:1:1 FGF23–FGFR1c–αKlotho–HS signal transduction complex.

Asymmetric FGFR dimerization is general

Because the αKlotho coreceptor does not directly participate in FGFR^{S} recruitment, we suggested that asymmetric mode of receptor dimerization revealed by our FGF23–FGFR–αKlotho–HS cryo-EM structures may also be relevant to paracrine FGFs. To test this conjecture, we studied the impacts of asymmetric FGFR^{P}–FGFR^{S} dimer interface mutations on the abilities of FGFR1c and FGFR2b to mediate paracrine FGF signalling. These two FGFR isoforms were chosen because of their overlapping and unique ligand-binding specificity/promiscuity profile (Extended Data Fig. 5d). FGFR1c responds to paracrine FGF1 (a pan-FGFR ligand) and FGF4, whereas FGFR2b mediates the actions of FGF1, FGF3, FGF7, FGF10 and FGF22. For FGFR2b studies, we generated L6 cell lines expressing either wild-type (FGFR2b^{WT}) or FGFR2b mutants (that is, FGFR2b^{E250A}, FGFR2b^{R255A}, FGFR2b^{I257A} and FGFR2b^{Y281A}) analogous to the FGFR1c mutants mentioned above. Both FGFR1c and FGFR2b mutants were impaired in their capacity to undergo ligand-induced tyrosine trans auto-phosphorylation, which was also reciprocated in reduced PLCγ1 and FRS2α phosphorylation (Extended Data Fig. 6).

Next, we carried out receptor complementation and heterodimerization assays as was done for FGF23. For FGFR2b complementation study, we established L6-FGFR2b$^{\Delta SLBS}$, L6-FGFR2b$^{\Delta PLBS}$ and L6-FGFR2b$^{\Delta SLBS}$ + FGFR2b$^{\Delta PLBS}$ cell lines corresponding to FGFR1c cell lines used for FGF23 study. In response to FGF1 or FGF4 stimulation, cells expressing

FGFR1c$^{\Delta PLBS}$ or FGFR1c$^{\Delta SLBS}$ alone failed to elicit any appreciable FGFR1c signalling whereas L6-FGFR1c$^{\Delta SLBS}$ + FGFR1c$^{\Delta PLBS}$ cells responded with robust FGFR1c activation and signalling (Fig. 5e,f). Likewise, FGF1, FGF3, FGF7 and FGF10 each induced FGFR activation and signalling only in L6-FGFR2b$^{\Delta SLBS}$ + FGFR2b$^{\Delta PLBS}$ co-expressors (Extended Data Fig. 7). Lastly, we studied the possibility of FGFR heterodimerization by paracrine FGFs by focusing on: (1) FGFR1c–FGFR4 and FGFR1b–FGFR2b heterodimerizations by FGF1; (2) FGFR1b–FGFR2b heterodimerization by FGF10; and (3) FGFR2b–FGFR3b heterodimerization by FGF3. For the FGFR1b–FGFR2b heterodimerization assay, we generated an L6 cell line expressing FGFR1b$^{\Delta SLBS}$, equivalent to FGFR2b$^{\Delta SLBS}$. For FGF2b–FGFR3b heterodimerization by FGF3, we used wild-type FGFR3b (FGFR3b^{WT}) as the ΔPLBS equivalent because this isoform naturally does not respond to FGF3. FGF1 failed to activate FGFR signalling in FGFR1c$^{\Delta PLBS}$ and FGFR4$^{\Delta SLBS}$ cell lines. However, strong FGFR signalling was seen in the FGFR1c$^{\Delta PLBS}$ + FGFR4$^{\Delta SLBS}$ co-expressing cell line in response to FGF1 stimulation (Fig. 5g,h). Likewise, both FGF1 and FGF10 induced robust FGFR activation/signalling only in the FGFR1b$^{\Delta SLBS}$ + FGFR2b$^{\Delta PLBS}$ co-expressing cell line but not in the L6-FGFR1b$^{\Delta SLBS}$ and L6-FGFR2b$^{\Delta PLBS}$ cell lines (Extended Data Fig. 8a–c). Lastly, FGF3 provoked signalling in the FGFR2b$^{\Delta SLBS}$ + FGFR3b^{WT} co-expressing cell line but not in the L6-FGFR2b$^{\Delta SLBS}$ and L6-FGFR3b^{WT} cells (Extended Data Fig. 8d,e). Based on these extensive cell-based data we conclude that paracrine FGF signalling is mediated via HS-induced asymmetric 1:2 FGF–FGFR dimers, reminiscent of the HS- and Klotho-induced endocrine 1:2 FGF23–FGFR dimer (Extended Data Fig. 9).

Concluding remarks

The 1:2:1:1 FGF23–FGFR–αKlotho–HS asymmetric quaternary complex structures and supporting biochemical data presented in this manuscript reveal how Klotho and HS glycosaminoglycan coreceptors act in concert to promote asymmetric 1:2 endocrine FGF–FGFR dimerization necessary for receptor activation and hence FGF hormone signalling. In this model, Klotho coreceptors tether FGF hormone and its primary receptor together and thus generate a stable endocrine FGF–FGFR^{P} complex in the context of

1:1:1 FGF–FGFR^{P}–Klotho ternary complexes. In doing so, Klotho coreceptors effectively offset the HS incompetency in stabilizing the endocrine FGF–FGFR^{P} complex (Extended Data Fig. 9). The stabilized endocrine FGF–FGFR^{P} complex is then assisted by a HS coreceptor to recruit a secondary FGFR^{S}. Importantly, the Klotho coreceptor dependency restricts the site of action of FGF hormones to Klotho expressing tissues/organs (that is, kidney and parathyroid glands in the case of FGF23). Consistent with a lack of a direct role of Klotho in FGFR^{S} recruitment, we show that the asymmetric mode of receptor dimerization is also applicable to paracrine FGFs. Unlike FGF hormones, paracrine FGFs have substantial affinity for HS and hence can rely solely on HS as an obligatory coreceptor to both stably bind FGFR^{P} and recruit FGFR^{S}. Despite sharing structural similarity, Klotho and HS-induced endocrine 1:2 FGF–FGFR dimers are thermodynamically inferior to HS-induced paracrine 1:2 FGF–FGFR dimers. Specifically, owing to the weak HS binding affinity of the FGF hormone, FGFR^{S} is bound less tightly in the endocrine 1:2 FGF–FGFR dimer, which accounts for the weaker receptor activation/signalling capacity of FGF hormones relative to paracrine FGFs as posited by our ‘threshold model’ for FGF signalling specificity[38]. Remarkably, the asymmetric receptor dimerization model is compatible with receptor heterodimerization, which is likely to serve as an additional mechanism to qualitatively and quantitatively fine-tune FGF signalling.

Methods

Expression constructs

The ligation-independent In-Fusion HD cloning kit (no. 639648, Clontech) was used to construct neomycin and hygromycin resistant pEF1α-IRES-neo and pEF1α-IRES-hygro lentiviral vectors encoding wild-type full-length human FGFR2c, FGFR3c and FGFR4 according to the protocol previously described for the human FGFR1c lentiviral expression construct[6]. The pET-30a-based bacterial expression construct for N-terminally his-tagged human FGF23 (residues Tyr25 to Ile251) was described previously[6]. pHLsec expression vectors encoding the extracellular ligand-binding region (that is, D2–D3 region) of FGFR3c (residues Asp142 to Arg365; FGFR3c^{ecto}) and

FGFR4 (residues Asp142 to Arg365; FGFR4^{ecto}) were made following the same strategy described previously for the human FGFR1c^{ecto} (ref. [6]). Single and multiple site mutations and truncations were introduced into expression constructs encoding the wild-type proteins using a Q5 site-directed mutagenesis kit (no. E0554S, New England Biolabs Inc.). The expression constructs were verified by restriction enzyme digestion and DNA sequencing.

Recombinant protein expression and purification

To produce minimally glycosylated ectodomains of FGFR1c, 3c and 4, N-acetylglucosaminyltransferase I (GnTI^{-}) deficient HEK293S cells were transiently transfected with respective pHLsec expression constructs via cationic polymer polyethyleneimine (PEI, no. 23966-1, Polysciences, Inc.) following a published protocol[40]. Then, 24 h posttransfection, tryptone N1 (TN1, catalogue no. 19553, Organotechnie) was added to the medium to promote protein expression. At day three, secreted FGFR ectodomains from 1 l of conditioned medium were captured on a heparin affinity HiTrap column (no. 17040703, GE Healthcare) and eluted with 20 column volume of salt gradient (0–2 M). Fractions containing FGFR^{ecto} proteins were pooled, concentrated to 5 ml and applied to a Superdex-75 gel filtration column (no. 28989333, GE Healthcare). FGFR^{ecto} proteins were eluted isocratically in 25 mM HEPES pH = 7.5 buffer containing 1 M NaCl. Wild-type and mutated FGF23 proteins were expressed in *E. coli* as inclusion bodies, refolded in vitro and purified to homogeneity using sequential cation exchange and SEC following an established protocol[6]. Secreted αKlotho^{ecto} was purified from conditioned media of a HEK293S GnTI^{-} cell line stably expressing the entire extracellular domain of human αKlotho (residues Met1 to Ser981; αKlotho^{ecto}) using a heparin affinity HiTrap column followed by SOUCRE Q anion and Superdex 200 column chromatography, as described previously[6].

Cryo-EM specimen preparation, image processing, model building and refinement

FGF23–FGFR1c–αKlotho–HS, FGF23–FGFR3c–αKlotho–HS or FGF23–FGFR4–αKlotho–HS quaternary complexes were prepared by mixing FGF23 with one of the FGFR ectodomains, αKlotho^{ecto} and a heparin dodecasaccharide 12 (HO12, Iduron Ltd) using a 1:1:1:1 molar ratio. The mixtures were concentrated to approximately 5 mg ml^{-1}, applied to a Superdex 200 column (no. 28989335, GE Healthcare) and eluted isocratically in 25 mM HEPES pH = 7.5 buffer containing 100 mM NaCl. Peak fractions were analysed by SDS–PAGE and top fractions containing the highest concentration and purity of the quaternary complex were used directly for grid preparation without further concentration to avoid protein aggregation. The final concentrations of FGF23–FGFR1c–αKlotho–HS, FGF23–FGFR3c–αKlotho–HS and FGF23–FGFR4–αKlotho–HS complexes for grid preparation were 1.5, 2.4 and 1.5 mg ml^{-1}, respectively.

To prepare the cryo-EM grids, 2–3 µl of purified protein complex at approximately1.5–2.5 mg ml^{-1} was applied to glow discharged gold grid (UltrAuFoil). The grid was then blotted for 1–2 s under 0 or 1 force at 100% humidity using a Mark IV Vitrobot (FEI) before plunging into liquid ethane. Micrographs of the FGF23–FGFR1c–αKlotho–HS and FGF23–FGFR4–αKlotho–HS complexes were acquired on a Talos Arctica microscope with K2 direct electron detector at ×36,000 magnification (corresponding to 1.096 Å per pixel). Accumulated doses used were 50.37 e^{-}/Å^{2} and 53.84 e^{-}/Å^{2} for FGF23–FGFR1c–αKlotho–HS and FGF23–FGFR4–αKlotho–HS, respectively. Micrographs of the FGF23–FGFR3c–αKlotho–HS complex were collected on a Titan Krios microscope equipped with a K2 direct electron detector and an energy filter. The magnification used was ×130,000, with a pixel size of 1.048 Å, and an accumulated dose of 72.44 e^{-}/Å^{2}. Leginon[41] was used to target the holes with 5–100 nm of ice thickness, resulting in 10,186, 6,409 and 16,602 micrographs being collected for each of three quaternary complexes.

WARP[42] was used for motion correction and contrast transfer function estimation for all three cryo-EM datasets. Micrographs with an overall resolution worse than 5.5 Å were excluded. The final number of micrographs used were 9,501, 5,164 and 15,049, respectively, yielding more than one million particles for each complex. Particle stacks were then imported to cryoSPARC[43] for two-dimensional classification, ab-initio

reconstruction with three or four models and three-dimensional classification. Finally, 1,497,967 (FGF23–FGFR1c–αKlotho–HS), 291,540 (FGF23–FGFR3c–αKlotho–HS) and 856,877 (FGF23–FGFR4–αKlotho–HS) particles were used for heterogeneous refinement with C1 symmetry, resulting in 2.74, 3.20 and 3.03 Å resolutions, respectively. Components/domains from FGF23–FGFR1c–αKlotho X-ray structures (PDB: 5W21) were manually docked into cryo-EM density maps using Chimera[44] and the rigid body was refined. Initial models were then adjusted in Coot[45] and real-space refined in Phenix[46]. Refinement and model statistics are shown in Extended Data Table 1. Representative cryo-EM images, two-dimensional class averages and three-dimensional maps of each quaternary complex are shown in Extended Data Fig. 1.

MD simulation

The cryo-EM structure of the FGF23–FGFR1c–αKlotho–HS quaternary complex was solvated in a water box of $16 \times 16 \times 16\ nm^3$. The CHARMM-GUI server was used to generate the configuration, topology[47,48,49,50] and the parameter files with the CHARMM36m force field[51]. In addition to protein molecules, the simulation system included about 138,073 water molecules, 393 sodium and 393 chloride ions (mimicking the 150 mM NaCl present in the protein buffer), resulting in a total of 439,298 atoms. A 300 ns all-atom MD simulation trajectory was generated using GROMACS 2021 (ref. [52]) at 303 K using a time step of 2 fs. The cubic periodic boundary condition was used during the simulations and the van der Waals interaction was switched off from 1 nm to 1.2 nm. The long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method. Energy minimization was carried out using the steepest descent algorithm, followed by a 0.4 ns constant particle number, volume and temperature (NVT) and a 20 ns constant particle number, pressure and temperature (NPT) equilibration simulation by gradually decreasing force restraints from 1000 kJ $mol^{-1}\ nm^{-2}$ to 400 kJ $mol^{-1}\ nm^{-2}$ (for the NVT stage) and 400 kJ $mol^{-1}\ nm^{-2}$ to 40 kJ $mol^{-1}\ nm^{-2}$ (for NPT stage). At the conclusion of the equilibration steps, all force restraints were removed and the MD simulation was performed in the NPT ensemble.

Generation of cell lines and FGFR signalling assay

HEK293T cells (verified by a morphology check under microscope, mycoplasma negative in 4′,6-diamidino-2-phenylindole (DAPI)) were used for lentiviral vector packaging and production of high-titre viral particles. An L6 myoblast cell line (no. GNR 4, National Collection of Authenticated Cell Cultures) was used as the host for stable expression of full-length (transmembrane) human FGFR1c, FGFR2c, FGFR3c, FGFR4 and mutants thereof. L6 cells endogenously express HSPGs but are devoid of FGFRs and αKlotho coreceptor and hence are naturally non-responsive to FGF23. However, via controlled ectopic co-expression of cognate FGFRs and αKlotho or exogenous supplementation with soluble αKlotho^{ecto}, these cells can respond to FGF23 stimulation. Accordingly, L6 cells are excellent hosts for reconstitution studies of FGF23 signalling in a physiological environment. Both HEK293T and L6 cells were maintained in Dulbecco’s Modified Eagle Medium (DMEM, no. C11995500BT, Gibco) supplemented with 10% fetal bovine serum (FBS, no. FSD500, ExCell Bio), 100 U ml^{-1} of Penicillin and 100 µg ml^{-1} Streptomycin (no. P1400, Solarbio). Viral packaging and generation of the recombinant lentivirus particles in HEK293T cells were carried out using published protocols[33]. For stable expression of individual wild-type or mutated FGFR cell lines, 2×10^5 L6 cells were plated in six-well cell culture dishes and infected with lentivirus particles encoding given FGFR in the presence of polybrene (5 µg ml^{-1}; no. sc-134220, Santa Cruz Biotechnology). Stable transfectants were selected using G418 (0.5 mg ml^{-1}, no. HY-17561, MedChemExpress) or hygromycin (8 µg ml^{-1}, no. HY-B0490, MedChemExpress). For the FGFR1c+αKlotho^{TM} co-expressing cell line, L6 cells stably expressing FGFR1c^{WT} (resistant to G418) were infected with lentiviral particles encoding wild-type or mutated transmembrane αKlotho (αKlotho^{TM}) and the co-expressing cells were selected using hygromycin (80 µg ml^{-1}, no. HY-B0490, MedChemExpress).

For cell stimulation studies, parental and stably transfected L6 cells were seeded in 12-well cell culture plates at a density of 1×10^5 cells per well and maintained for 24 h. On the next day, the cells were rinsed three times with phosphate buffered saline (PBS) and then serum starved for 12 h and

costimulated with FGF23^{WT} and αKlotho^{ecto} for 5, 10, 20 and 40 min. For single time point stimulation, samples were harvested after 5 min.

After stimulation, cells were lysed and total lysates samples were analysed by western blotting, as previously described[33]. The following antibodies were used: phosphorylated FGFR (1:1000, no. 3471S, Cell Signaling Technology), phosphorylated FRS2α (1:1000, no. 3864S, Cell Signaling Technology), phosphorylated PLCγ1 (1:1000, no. 2821S, Cell Signaling Technology), phosphorylated ERK1/2 (1:1000, no. 4370S, Cell Signaling Technology), α-tubulin (1:20000, no. 66031-1-Ig, Proteintech), total-FGFR1 (1:1000, no. 9740S, Cell Signaling Technology), total-FGFR2 (1:1000, no. 23328S, Cell Signaling Technology), total-FGFR3 (1:1000, no. ab133644, Abcam), total-FGFR4 (1:1000, no. 8562S, Cell Signaling Technology), HRP conjugated goat anti-mouse IgG (H + L) (1:5000, no. SA00001-1, Proteintech), HRP conjugated goat anti-rabbit IgG(H + L) (1:5000, no. SA00001-2, Proteintech). Blots were developed using enhanced chemiluminescence reagents (no. P10300, NCM Biotech Laboratories) by the ChemiDoc XRS+ system (Bio-Rad) or Amersham ImageQuant 800 (GE).

Determination of cell-surface expression of mutated FGFRs via endoglycosidase H sensitivity assay

Endoglycosidase H (Endo H) sensitivity was used to analyse potential impacts of ectodomain mutations on FGFR glycosylation/maturation and hence trafficking to the cell surface. In immunoblots, wild-type FGFRs migrate as a doublet of a major diffuse upper and a minor sharp lower band. The upper band represents the fully glycosylated mature FGFR, decorated with complex sugars that has passed the ER quality control and has been successfully trafficked to the cell surface. On the other hand, the faster migrating lower band is an incompletely processed high mannose form that is trapped in ER. Mutations affecting receptor maturation manifest in an increase in proportion of the faster migrating ER-resident band. The mannose-rich form is sensitive to Endo H which cleaves the bond between two N-acetylglucosamine (GlcNAc) subunits directly proximal to the asparagine residue. However, the fully glycosylated cell surface-resident band is resistant to Endo H. Accordingly, cell-surface abundance of FGFRs

can be expressed as a ratio of Endo H-resistant fraction over the total receptor expression as determined by treating the receptor with PNGase F. This enzyme is an amidase that hydrolyzes the bond between the innermost GlcNAc and asparagine irrespective of complex sugar content and thus completely strips the FGFR from all its N-linked sugars. Accordingly, WT or mutant FGFR cell lines were lysed in an NP-40 lysis buffer (Biotime, no. P0013F, supplemented with 1 mM PMSF) for 15 min at 4 °C. First, 20 μg total protein (quantified by BCA assay) were denatured with glycoprotein denaturing buffer at 100 °C for 10 min and then treated with 500 units of Endo H (New England Biolabs, no. P0702S) or peptide-N-glycosidase F (PNGase F) (New England Biolabs, no. P0704S) following the manufacturer's instructions. Endo H- and PNGase F-treated samples were immunoblotted with FGFR isoform-specific antibodies, as detailed above.

Size-exclusion chromatography–multi-angle dynamic light scattering

The molecular mass of the SEC-purified FGF23–FGFR1c-αKlotho–HS quaternary complex was determined by multi-angle light scattering following the established protocol[6]. Before the experiment, at least 60 ml of degassed running buffer (25 mM HEPES pH 7.5 containing 150 mM NaCl) were passed through the system to equilibrate the column and establish steady baselines for light scattering and refractive index detectors. Then, 50 μl of purified FGF23–FGFR1c-αKlotho–HS quaternary complex (1.5 mg ml^{-1}) was injected onto the Superdex 200 10/300 GL column and the eluent was continuously monitored at 280 nm absorbance, laser light scattering and refractive index at a flow rate of 0.5 ml min^{-1}. As a control, 50 μl of a purified FGF23–FGFR-αKlotho ternary complex (1.5 mg ml^{-1}) sample was analysed under the same condition. The experiments were performed at ambient temperature. Laser light scattering intensity and eluent refractive index values were used to derive molecular mass as implemented by the ASTRA software (Wyatt Technology Corp.).

Proximity ligation assay

Cells were seeded onto microscope cover glasses (no. WHB-12-CS-LC, WHB) placed inside 12-well cell culture dishes at 1×10^5 cells per well and

allowed to adhere for 24 h. On the next day, cells were washed three times with PBS and serum starved for 12 h. Following costimulation with $FGF23^{WT}$ and $\alpha Klotho^{ecto}$ for 20 min, cells were washed three times with PBS and fixed with 4% paraformaldehyde for 30 min prior to PLA. The PLA reaction was performed using a Duolink PLA kit following the manufacturer's instructions (no. DUO92101, Sigma-Aldrich) and visualized via fluorescence microscopy. Briefly, slides were treated with blocking solution for 60 min at 37 °C, rinsed three times with wash buffer A and then incubated overnight at 4 °C with two different primary antibodies raised in two different species for each FGFR isoform of interest (FGFR1 (1:20, no. PA5-25979, ThermoFisher) from rabbit, FGFR1 (1:100, no. ab824, Abcam) from mouse, FGFR2 (1:100, no. 23328S, Cell Signaling Technology) from rabbit, FGFR2 (1:50, no. sc-6930, Santa Cruz Biotechnology) from mouse, FGFR3 (1:100, no. MA5-32620, ThermoFisher) from rabbit, FGFR3 (1:100, no. sc-13121, Santa Cruz Biotechnology) from mouse, FGFR4 (1:100, no. 8562S, Cell Signaling Technology) from rabbit, FGFR4 (1:100, no. sc-136988, Santa Cruz Biotechnology) from mouse). Following three rinses with wash buffer A, slides were incubated with oligo-linked secondary antibodies (Duolink anti-mouse minus and anti-rabbit plus) for 1 h at 37 °C. Slides were rinsed again with wash buffer A and immersed in ligase solution for 30 min at 37 °C, to allow formation of circular DNA, followed by incubation with polymerase solution for 100 min at 37 °C for rolling circle amplification in a dark room. Slides were rinsed twice with wash buffer B for 10 min each, followed by rinsing with a 100-fold dilution of buffer B for 1 min. For imaging analysis, slides were covered by coverslips using a minimal volume of Duolink PLA mounting medium containing DAPI. Slides were examined using a confocal laser scanning microscope (C2si, Nikon).

Statistical analysis and reproducibility

All statistical analyses were carried out using GraphPad Prism 8.0. For statistical analysis of immunoblotting data, densitometric values (determined using ImageJ) from three independent experiments were used. For statistical analysis of PLA data, a number of fluorescent dots and cells (counted manually) was used from six randomly chosen microscope fields from two biologically independent experiments. Processing of the western blotting

and PLA data in Figs. 2c,d, 3b–d and 4b,c was done using two-way ANOVA, followed by Tukey. One-way ANOVA, followed by Tukey was applied to process the western blotting data in Extended Data Fig. 3. Protein purifications were repeated at least eight times yielding samples with comparable purity/quantity. Western blotting experiments were carried out in biological triplicates with similar results. PLA assays were repeated at least twice independently, with analogous results.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Electron density maps and refined models for the FGF23–FGFR1c–αKlotho–HS (EMD-34075, PDB: 7YSH), FGF23–FGFR3c–αKlotho–HS (EMD-34082, PDB: 7YSU) and FGF23–FGFR4–αKlotho–HS (EMD-34084, PDB: 7YSW) quaternary complexes have been deposited in the Electron Microscopy Data Bank and will be released upon acceptance of the manuscript. Raw uncropped western blot images are compiled in Supplementary Fig. 1. Source data are provided with this paper.

References

1. Kliewer, S. A. & Mangelsdorf, D. J. A dozen years of discovery: insights into the physiology and pharmacology of FGF21. *Cell Metab.* **29**, 246–253 (2019).

2. Degirolamo, C., Sabba, C. & Moschetta, A. Therapeutic potential of the endocrine fibroblast growth factors FGF19, FGF21 and FGF23. *Nat. Rev. Drug Discov.* **15**, 51–69 (2016).

3. Kuro-o, M. Klotho and betaKlotho. *Adv Exp Med Biol* **728**, 25–40 (2012).

4. Razzaque, M. S. The FGF23-Klotho axis: endocrine regulation of phosphate homeostasis. *Nat. Rev. Endocrinol.* **5**, 611–619 (2009).

5. Kuro, O. M. The Klotho proteins in health and disease. *Nat. Rev. Nephrol.* **15**, 27–44 (2019).

6. Chen, G. et al. alpha-Klotho is a non-enzymatic molecular scaffold for FGF23 hormone signalling. *Nature* **553**, 461–466 (2018).

7. Itoh, N. & Ornitz, D. M. Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease. *J. Biochem.* **149**, 121–130 (2011).

8. Itoh, N. & Ornitz, D. M. Evolution of the Fgf and Fgfr gene families. *Trends Genet.* **20**, 563–569 (2004).

9. Kharitonenkov, A. et al. FGF-21 as a novel metabolic regulator. *J. Clin. Invest.* **115**, 1627–1635 (2005).

10. Cariello, M., Piglionica, M., Gadaleta, R. M. & Moschetta, A. The enterokine fibroblast growth factor 15/19 in bile acid metabolism. *Handb. Exp. Pharmacol.* **256**, 73–93 (2019).

11. Wang, X. & Shapiro, J. I. Evolving concepts in the pathogenesis of uraemic cardiomyopathy. *Nat. Rev. Nephrol.* **15**, 159–175 (2019).

12. Reitman, M. L. FGF21 mimetic shows therapeutic promise. *Cell Metab.* **18**, 307–309 (2013).

13. Zhao, Y., Dunbar, J. D. & Kharitonenkov, A. FGF21 as a therapeutic reagent. *Adv. Exp. Med. Biol.* **728**, 214–228 (2012).

14. Kharitonenkov, A. & DiMarchi, R. FGF21 revolutions: recent advances illuminating FGF21 biology and medicinal properties. *Trends Endocrinol. Metab.* **26**, 608–617 (2015).

15. Gaich, G. et al. The effects of LY2405319, an FGF21 analog, in obese human subjects with type 2 diabetes. *Cell Metab.* **18**, 333–340 (2013).

16. Samuel, V. T. & Shulman, G. I. Nonalcoholic fatty liver disease as a nexus of metabolic and hepatic diseases. *Cell Metab.* **27**, 22–41 (2018).

17. Harrison, S. A. et al. NGM282 for treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **391**, 1174–1185 (2018).

18. Dickson, I. NASH: Successful phase II trial of FGF19 analogue. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 256 (2018).

19. Johnson, K. et al. Therapeutic effects of FGF23 c-tail Fc in a murine preclinical model of X-linked hypophosphatemia via the selective modulation of phosphate reabsorption. *J. Bone Miner. Res.* **32**, 2062–2073 (2017).

20. Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117–1134 (2010).

21. Katoh, M. Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat. Rev. Clin. Oncol.* **16**, 105–122 (2019).

22. Yeh, B. K. et al. Structural basis by which alternative splicing confers specificity in fibroblast growth factor receptors. *Proc. Natl Acad. Sci. USA* **100**, 2266–2271 (2003).

23. Olsen, S. K. et al. Structural basis by which alternative splicing modulates the organizer activity of FGF8 in the brain. *Genes Dev* **20**, 185–198 (2006).

24. Liu, Y. et al. Regulation of receptor binding specificity of FGF9 by an autoinhibitory homodimerization. *Structure* **25**, 1325–1336 e1323 (2017).

25. Sarrazin, S., Lamanna, W. C. & Esko, J. D. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* **3**, a004952 (2011).

26. Schlessinger, J. et al. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* **6**, 743–750 (2000).

27. Chen, L. et al. Molecular basis for receptor tyrosine kinase A-loop tyrosine transphosphorylation. *Nat. Chem. Biol.* **16**, 267–277 (2020).

28. Goetz, R. et al. Molecular insights into the klotho-dependent, endocrine mode of action of fibroblast growth factor 19 subfamily members. *Mol. Cell. Biol.* **27**, 3417–3428 (2007).

29. Yu, X. et al. Analysis of the biochemical mechanisms for the endocrine actions of fibroblast growth factor-23. *Endocrinology* **146**, 4647–4656 (2005).

30. Urakawa, I. et al. Klotho converts canonical FGF receptor into a specific receptor for FGF23. *Nature* **444**, 770–774 (2006).

31. Ito, S. et al. Impaired negative feedback suppression of bile acid synthesis in mice lacking betaKlotho. *J. Clin. Invest.* **115**, 2202–2208 (2005).

32. Adams, A. C., Cheng, C. C., Coskun, T. & Kharitonenkov, A. FGF21 requires betaklotho to act in vivo. *PLoS ONE* **7**, e49977 (2012).

33. Kharitonenkov, A. et al. FGF-21/FGF-21 receptor interaction and activation is determined by betaKlotho. *J. Cell. Physiol.* **215**, 1–7 (2008).

34. Kurosu, H. et al. Tissue-specific expression of betaKlotho and fibroblast growth factor (FGF) receptor isoforms determines metabolic activity of FGF19 and FGF21. *J. Biol. Chem.* **282**, 26687–26695 (2007).

35. Goetz, R. et al. Klotho coreceptors inhibit signaling by paracrine fibroblast growth factor 8 subfamily ligands. *Mol. Cell. Biol.* **32**, 1944–1954 (2012).

36. Imura, A. et al. Secreted Klotho protein in sera and CSF: implication for post-translational cleavage in release of Klotho protein from cell membrane. *FEBS Lett.* **565**, 143–147 (2004).

37. Bloch, L. et al. Klotho is a substrate for alpha-, beta- and gamma-secretase. *FEBS Lett.* **583**, 3221–3224 (2009).

38. Zinkle, A. & Mohammadi, M. A threshold model for receptor tyrosine kinase signaling specificity and cell fate determination. *F1000Res.* **7**, 872 (2018).

39. Patstone, G. & Maher, P. Copper and calcium binding motifs in the extracellular domains of fibroblast growth factor receptors. *J. Biol. Chem.* **271**, 3343–3346 (1996).

40. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243–1250 (2006).

41. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).

42. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).

43. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

44. Pettersen, E. F. et al. UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

46. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr*. **66**, 213–221 (2010).

47. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

48. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **97**, 50–58 (2009).

49. Wu, E. L. et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).

50. Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

51. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).

52. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).

Acknowledgements

Financial support was provided by the National Key R&D Program of China (2017YFA0506000 to X.L.), National Natural Science Foundation of China (82073705 & 82273842 to G.C., 82103999 to L.C., 81930108 to G.L.), Oujiang laboratory startup fund (OJQD2022007 to M.M.), Kungpeng Action Plan award (to M.M.), Natural Science Funding of Zhejiang Province (LR22H300002 to G.C., LQ22H300007 to L.C.), Wenzhou Major Scientific and Technological Innovation Project (ZY2021023 to G.C.), Qianjiang Talent Plan of Zhejiang (QJD1902016 to G.C.) and Wenzhou Institute (UCAS) startup fund (WIUCASQD2021043, to Y.W.).

Author information

Author notes

1. These authors contributed equally: Lingfeng Chen, Lili Fu, Jingchuan Sun, Zhiqiang Huang, Mingzhen Fang

Authors and Affiliations

1. Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine, Vision, and Brain Health), School of Pharmaceutical Sciences, Wenzhou Medical University, Wenzhou, China

   Lingfeng Chen, Lili Fu, Jingchuan Sun, Zhiqiang Huang, Mingzhen Fang, Xin Liu, Junliang Lu, Zixiang Pan, Yang Wang, Xiaokun Li, Gaozhi Chen & Moosa Mohammadi

2. School of Pharmaceutical Sciences, Hangzhou Medical College, Hangzhou, China

   Lingfeng Chen & Guang Liang

3. Institute of Cell Growth Factor, Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine, Vision, and Brain Health), Wenzhou, China

   Lili Fu, Zhiqiang Huang, Mingzhen Fang, Xin Liu, Junliang Lu, Zixiang Pan, Gaozhi Chen & Moosa Mohammadi

4. State Key Laboratory for Macromolecule Drugs and Large-scale Preparation, Wenzhou Medical University, Wenzhou, China

   Lili Fu & Xiaokun Li

5. Laboratory of Cell Fate Control, School of Life Sciences, Westlake University, Hangzhou, China

   Jingchuan Sun

6. Department of Physiology and Cellular Biophysics, Columbia University Irving Medical Center, New York, NY, USA

   Allen Zinkle

7. Center of Biomedical Physics, Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, China

   Yang Wang

8. National Engineering Research Center of Cell Growth Factor Drugs and Protein Biologics, Wenzhou Medical University, Wenzhou, China

   Xiaokun Li

9. Institute of chronic kidney disease, Wenzhou Medical University, Wenzhou, China

   Gaozhi Chen

Contributions

L.C. expressed, purified and prepared the quaternary complexes, carried out the size-exclusion chromatography–multi-angle light scattering analysis and prepared the structural figures. L.C. and J.S. prepared the cryo-EM grids, and collected and processed the images. A.Z. expressed and purified FGFR ectodomains. L.F. generated the mammalian expression constructs. J.L. expressed the mutant FGF23 proteins. L.F., Z.H., M.F., X.L., Z.P. and G.C. generated western blotting and PLA data and prepared the related figures. Y.W. generated MD simulation data. M.M. conceived the project, built and analysed the structural models and wrote the manuscript. All authors participated in discussion and revision of the manuscript.

Corresponding authors

Correspondence to Xiaokun Li, Gaozhi Chen or Moosa Mohammadi.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Makoto Kuro-o and the other, anonymous, reviewers(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Determination of molecular structures of FGF23-FGFR-αKlotho-HS quaternary complexes by Cryo-EM.

Image processing workflow for the FGF23-FGFR-αKlotho-HS quaternary complex containing FGFR1c (left), FGFR3c (middle) or FGFR4 (right) as the receptor component. Gold-standard Fourier shell correlation (GSFSC) show global resolutions of 2.74, 3.20, and 3.03 Å for FGF23-FGFR1c-αKlotho-HS, FGF23-FGFR3c-αKlotho-HS, and FGF23-FGFR4-αKlotho-HS quaternary complexes, respectively.

Extended Data Fig. 2 Overall topology of the FGF23 signaling complex and its strict αKlotho dependency.

A, Cartoon representation of the cryo-EM structure of the 1:2:1:1 FGF23-FGFR-αKlotho-HS quaternary complex in four different orientations related by 90° rotation along the vertical axis. The asymmetric quaternary complex has an average dimension of 130 Å × 100 Å × 55 Å. The proximity of membrane insertion points of FGFR chains (~25 Å apart) would be conducive to the formation of an asymmetric A-loop trans-phosphorylating dimer of intracellular kinase domain[27]. FGF23, αKlotho and HS are shown

in orange, blue and magenta, respectively. Primary receptor ($FGFR^{P}$) and secondary receptor ($FGFR^{S}$) are shown in pale green and cyan. Dashed lines denote residues 172–182 of FGF23, the linker between FGF23's trefoil core and its distal αKlotho binding site, which could not be built due to lack of interpretable electron density. This region does not interact with either αKlotho or FGFR and is likely disordered/flexible. Notably, this region harbors the regulatory subtilisin-like proprotein convertase (SPC) site,^{176}RHT^{178}R^{179}/S^{180}AE^{182}, which includes a furin type protease cleavage site (R179), an O-glycosylation site (T178) and a serine phosphorylation site (S180). Three enzymes namely GalNAc-T3(N-acetylgalactosaminyltransferase3), Fam20C (the family with sequence similarity 20, member C), and a yet to be discovered furin type protease converge on this site to regulate FGF23 processing. The high flexibility of this region is likely necessary for the action of these enzymes. **B**, FGF23 signaling is strictly αKlotho dependent. L6-$FGFR1c^{WT}$ cells were treated with increasing concentrations of recombinant $FGF23^{WT}$ alone, in combination with soluble αKlotho or left untreated. Whole cell lysates immunoblotted as in Fig. 2c. Note that even supra pharmacological concentrations (as high as 10 micromolar) of $FGF23^{WT}$ fails to activate FGFR1c signaling. However, when co-treated with soluble αKlotho, as little as 10 nM FGF23 induces a robust FGFR1c activation. Experiments were performed in biological triplicates with similar results. **C**, Superimposition of X-ray structure of the FGF23–FGFR1c–αKlotho ternary complex (PDB ID: 5W21, colored in green) onto the corresponding FGF23–$FGFR1c^{P}$–αKlotho portion within the cryo-EM structure of 1:2:1:1 FGF23-FGFR1c-αKlotho-HS quaternary complex (colored in pink) shows high degree of similarity (RM)D of 1.16 Å].

Extended Data Fig. 3 Cell surface expression analysis of mutated FGFRs via Endoglycosidase H (Endo H) sensitivity assay.

a, Denatured lysates from L6-$FGFR1c^{WT}$, L6-$FGFR1c^{K175Q/K177Q}$ ($FGFR1c^{\Delta HBS1}$), L6-$FGFR1c^{K207Q/R209Q}$ ($FGFR1c^{\Delta HBS2}$), and L6-$FGFR1c^{K175Q/K177Q/K207Q/R209Q}$ ($FGFR1c^{\Delta HBS1+2}$) cells were incubated with Endoglycosidase H (Endo H) or Peptide-N-Glycosidase F (PNGase F) or left

untreated. Samples were immunoblotted with FGFR1c isoform-specific antibody. Immunoblotting data were quantitated as described in the Methods section and are presented as the mean ± SD. **b**–**d**, Denatured lysates from L6 cell lines stably expressing $FGFR1c^{WT}$, its four *Site 2* ($FGFR1c^{E249A}$, $FGFR1c^{R254A}$, $FGFR1c^{I256A}$, $FGFR1c^{Y280A}$) and one *Site 1* ($FGFR1c^{A170D/A171D/S219D}$) mutants (**b**), wild-type FGFR3c ($FGFR3c^{WT}$) or its four *Site 2* mutants ($FGFR3c^{E247A}$, $FGFR3c^{R252A}$, $FGFR3c^{I254A}$, $FGFR3c^{Y278A}$) (**c**), wild-type FGFR4 ($FGFR4^{WT}$) or its four *Site 2* mutants (E243A, R248A, I250A, Y274A) (**d**) were treated with Endo H, PNGase F or left untreated. Samples were immunoblotted with FGFR isoform-specific antibodies as indicated. Experiments were performed in biological triplicates with similar results. Quantitation was done as described in the Methods section and are presented as the mean ± S.D. *P* values were determined by One-way ANOVA followed by Tukey's multiple comparisons *post hoc* test.

Source data

Extended Data Fig. 4 FGF23 N-terminus experiences a major conformational change in the quaternary complex.

A, Left and Center: Cartoon representations of the FGF23-$FGFR^S$ and $FGFR^P$-αKlotho components of the quaternary complex shown in the orientation obtained via alignment of D2 domains of $FGFR^S$ and $FGFR^P$. Right: Superimposition of FGF23-$FGFR^S$ and $FGFR^P$-αKlotho components via aligning receptors' D3 domains. Note that FGF23 N-terminus and αKlotho RBA engage the equivalent hydrophobic grooves in D3 of $FGFR^S$ and $FGFR^P$ chains. **b**–**d**, FGF23 N-terminus undergoes a major conformational change in FGF23–FGFR–αKlotho–HS quaternary complexes. Cartoon and surface (only for FGF23 component) representation of HS-free ternary complex (that is, X-ray structure) (**b**) and HS-induced quaternary complex (that is, cryo-EM structures) (**c**) in the same orientation obtained via alignment of their FGF23 chains. (**d**) Cartoon representation of an overlay between FGF23-FGFR1c from FGF23–FGFR1c–αKlotho ternary complex and FGF23-$FGFR1c^P$ from quaternary complex via FGF23 alignment. Note the dramatic difference in the positions of FGF23 N-terminus between the two structures. **e**, Electron densities of FGF23 N-

terminus in cryo-EM structures of FGF23–FGFR1c–αKlotho–HS quaternary complexes at the indicated contour levels. Note that the electron densities for FGF23 N-terminus are weak and patchy. f–i, MD simulation data show that FGF23 N-terminus engages D3 of FGFR1c^{S}. f, Left: Cartoon representation of the cryo-EM structure of FGF23-FGFR1c-αKlotho-HS quaternary complex. FGF23, FGFR1c^{P}, FGFR1c^{S} and αKlotho are in orange, green, cyan, and blue, respectively. Right: Zoomed-in view of the boxed region (left) showing conformational progression of the N-terminal tail of FGF23 (that is, Y25 to W36) and D3 of FGFR1c^{S} in a 300 ns MD simulation trajectory denoted by a color transition from the orange (0 ns) to blue (300 ns) in 60 ns intervals. FGF23 N-terminus interacts with the three stranded βC: βF: βG sheet in FGFR1c^{S} D3 domain. **g**-**h**, Changes in RMSD (**g**) and Root Mean Square Fluctuation (RMSF, **h**), respectively, of N-terminal tail of FGF23 during MD simulation. Note that RMSD of N-terminal residues of FGF23 stabilized around 4 Å after 120 ns. Importantly, residues at the distal and proximal ends of FGF23 N-terminus exhibited largest and smallest RMSF, respectively, mirroring their respective cryo-EM electron densities (compare e and h). **i**, Changes in the distances of four selected contact pairs (Y25–L342^{S}, S29–R254^{S}, L31–A259^{S} and L32–I256^{S}) between N-terminal residues of FGF23 and residues in D3 of FGFR1c^{S}. The distances of Y25–L342^{S}, L31–A259^{S}, and L32–I256^{S} hydrophobic residue pairs fluctuated around 5 Å indicative of formation of hydrophobic contacts between these residue pairs. Likewise, the pairwise distance for S29-R254^{S} fluctuated around 4 Å after 120 ns indicative of hydrogen bonding between side chains of this residue pair.

Extended Data Fig. 5 Rationale for cell-based receptor complementation assay and structure-based sequence alignment of seven FGFR isoforms.

a-**b**, Rationale for receptor complementation assay. Center: Cartoon/surface representation of the FGF23–FGFR1c–αKlotho–HS quaternary complex in two orientations related by 180° rotation around Y axis. Circles denote approximate circumferences of ligand binding sites in primary and secondary receptors selected for mutagenesis. **a**, Left: expanded view of circled region on the secondary receptor showing that Ile-203 and Ser-223 of

FGFR1c^{S} engage in hydrophobic and hydrogen bonding contacts with residues in FGF23's core. Right: expanded view of circled region on the primary receptor showing that Ala-167 and Val-248 of FGFR1c^{P} engage in highly conserved hydrophobic contacts with Tyr-124 and Leu-158 of FGF23, respectively. Additionally, Ala-167 also makes a hydrogen bond with Tyr-124. **b**, Left, expanded view of the circled region on the primary receptor showing that the corresponding Ile-203 and Ser-223 in FGFR1c^{P} are solvent exposed. Thus, an engineered FGFR1c$^{I203E/S223E}$ molecule (i.e., FGFR1c$^{\Delta SLBS}$) should preserve the ability to act as primary receptor despite losing the capacity to function as secondary receptor. Right, expanded view of the circled region on secondary receptor showing that the corresponding Ala-167 and Val-248 in FGFR1c^{S} do not play any role in the quaternary complex formation and are solvent exposed. Thus, an engineered FGFR1c$^{A167D/V248D}$ mutant (i.e., FGFR1c$^{\Delta PLBS}$) is predicted to lose the ability to function as primary receptor but could still act as secondary receptor. **c**, Demonstration of receptor complementation between FGFR1c$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ and between FGFR4$^{\Delta SLBS}$ and FGFR1c$^{\Delta PLBS}$ via PLA. Representative fluorescent microscopy fields of stable cell lines subjected to PLA are shown. Scale bar,10 μm. Experiments were performed at least two times biological repeat with similar results. **d**, Structure-based sequence alignment of ligand binding regions (i.e., D2, D2-D3 linker and D3) of all seven FGFR isoforms. As a guide, a schematic representation of a prototypical FGFR is shown and its various domains are labeled. The C-terminal half of D3 which undergoes alternative splicing in FGFR1-FGFR3 is highlighted in dark blue. Note that the unspliced N-terminal half of D3 is conserved between b and c isoforms. Secondary structure elements are provided on the top of alignment. The HS interacting residues are all localized within D2 domain and are colored in magenta. Residues mediating the FGF23-FGFR^{P} and αKlotho-FGFR^{P} interfaces are colored yellow and green, respectively. Note that residues mediating direct FGFR^{P}-FGFR^{S} contacts (in light blue) are fully conserved amongst all seven isoforms. As for the FGF23-FGFR^{S} interface (in orange), only D2 residues that contact FGF23's core, are conserved between the seven isoforms. However, two hydrophobic D3 residues that interact with FGF23 N-terminus at the FGF23-FGFR^{S} interface are conserved only in FGFR1c-3c and FGFR4 (indicated by black arrow heads). In FGFR1b-3b, these residues,

which overlap with binding site for RBA of αKlotho, are replaced by charged residues. Hydrophobicity of D3 groove FGFR1b-3b is further disrupted by the presence of a unique N-linked glycosylation site within their D3 grooves (denoted by a red box). In FGFR2c, a key conserved hydrophobic residue is replaced by a charged Lys-296 (underlined in red) which likely accounts for the inability of FGFR2c to bind αKlotho.

Extended Data Fig. 6 The asymmetric $FGFR^{P}$-$FGFR^{S}$ interface observed in FGF23-FGFR-αKlotho-HS cryo-EM structures is required for multiple paracrine FGF signaling.

a, Immunoblots of whole cell extracts from untreated or FGF-treated (1 nM) untransfected and transfected L6 cell lines stably expressing $FGFR1c^{WT}$, $FGFR1c^{E249A}$, $FGFR1c^{R254A}$, $FGFR1c^{I256A}$, or $FGFR1c^{Y280A}$ probed with antibodies against phosphorylated FGFR, PLCγ1 and FRS2α. **b**, Immunoblots of whole cell extracts from untreated or FGF-treated (1 nM for FGF1 and 2 nM for FGF3/7/10/22) treated untransfectedand transfected L6 cell lines stably expressing $FGFR2b^{WT}$, $FGFR2b^{E250A}$, $FGFR2b^{R255A}$, $FGFR2b^{I257A}$, or $FGFR2b^{Y281A}$ probed with FGFR2b isoform specific antibody (top) or phosphospecific antibodies as in panel **a**. Experiments were performed in biological triplicates with similar results.

Extended Data Fig. 7 Receptor complementation assays demonstrate the asymmetry of four distinct paracrine FGF-FGFR2b-HS signaling complexes.

a, Schematic diagram showing that in response to paracrine FGF1/3/7/10 and HS, $FGFR2b^{\Delta SLBS}$ and $FGFR2b^{\Delta PLBS}$ can complement each other and form 1:1:1:1 FGF-$FGFR2b^{\Delta SLBS}$-$FGFR2b^{\Delta PLBS}$-HS asymmetric signaling complexes. **b**-**e**, Demonstration of receptor complementation between $FGFR2b^{\Delta SLBS}$ and $FGFR2b^{\Delta PLBS}$ via westernblotting. L6 cell lines singly expressing $FGFR2b^{WT}$, $FGFR2b^{\Delta PLBS}$, $FGFR2b^{\Delta SLBS}$, or co-expressing $FGFR2b^{\Delta SLBS}$ with $FGFR2b^{\Delta PLBS}$ were treated with 1 nM FGF1 (**b**), 2 nM FGF3 (**c**), 2 nM FGF7 (**d**), or 2 nM FGF10 (**e**) for increasing time intervals,

and total cell extracts were immunoblotted as indicated. Experiments were performed in biological triplicates with similar results.

Extended Data Fig. 8 Receptor heterodimerization assays validate the asymmetry of paracrine FGF-FGFR-HS signaling complexes.

a, Schematic diagram showing that FGF1 or FGF10, HS, and FGFR1b$^{\Delta SLBS}$ (serving as primary receptor) form stable complexes which subsequently recruit FGFR2b$^{\Delta PLBS}$ as secondary receptor. **b**-**c**, L6 cell lines singly expressing FGFR1b^{WT} or FGFR1b$^{\Delta SLBS}$, or co-expressing FGFR1b$^{\Delta SLBS}$ + FGFR2b$^{\Delta PLBS}$ were treated with 1 nM FGF1 (**b**) or 2 nM FGF10 (**c**) for increasing time intervals and cell extracts were immunoblotted. **d**, Schematic diagram showing that in the presence of HS, FGF3 and FGFR2b$^{\Delta SLBS}$ (serving as primary receptor) form a stable complex and subsequently recruit FGFR3b^{WT} as secondary receptor. **e**, L6 cell lines expressing FGFR3b^{WT} alone or co-expressing it with FGFR2b$^{\Delta SLBS}$ were treated with 2 nM FGF3 for increasing time intervals and cell extracts were immunoblotted. Experiments were performed in biological triplicates with similar results.

Extended Data Fig. 9 Asymmetric receptor dimerization is a universal mechanism in FGF signaling.

a, Due to FGF hormone's weak HS binding affinity, HS alone is incompetent in stabilizing the endocrine FGF-FGFR complex and inducing sustained asymmetric receptor dimerization/activation. Blurring and loose association are used to emphasize the unstable/transient nature of putative FGF-FGFR-HS ternary complex and physiologically inconsequential receptor dimerization/activation. **b**, Membrane bound Klotho co-receptor simultaneously engages FGFR's D3 domain and FGF's C-terminal tail thereby stabilizing the endocrine FGF-FGFR complex within a ternary complex. In so doing, Klotho co-receptor effectively compensates for HS incompetency in stabilizing binary endocrine FGF-FGFR complex. HS is now in position to recruit a second FGFR to the stabilized binary complex thus inducing asymmetric dimerization. Nevertheless, due to FGF hormone's

weak HS binding affinity, Klotho and HS-induced endocrine 1:2 FGF-FGFR dimers are still inferior to HS-induced paracrine 1:2 FGF-FGFR dimers in terms of longevity/stability (indicated by slight blurring of $FGFR^S$ in **b**). **c**, Due to their high HS binding affinities, paracrine FGFs can rely on HS as the sole co-receptor to stably bind primary FGFR and recruit a secondary FGFR, thereby inducing formation of rigid and long-lived asymmetric receptor dimers.

Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics

Supplementary information

Supplementary Information

Supplementary Figs. 1–4 and Tables 1–2.

Reporting Summary

Peer Review File

Source data

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 3

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation,

distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Chen, L., Fu, L., Sun, J. *et al.* Structural basis for FGF hormone signalling. *Nature* **618**, 862–870 (2023). https://doi.org/10.1038/s41586-023-06155-9

- Received: 06 September 2022
- Accepted: 02 May 2023
- Published: 07 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06155-9

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06155-9

Amendments & Corrections

- **Author Correction: Single-neuronal predictions of others' beliefs in humans** [01 Jun 2023]
  -
- **Publisher Correction: Crop switching can enhance environmental sustainability and farmer incomes in China** [02 Jun 2023]
  -
- **Publisher Correction: Targeting PD-L2–RGMb overcomes microbiome-related immunotherapy resistance** [01 Jun 2023]
  -
- **Publisher Correction: Cortico-cortical feedback engages active dendrites in visual cortex** [06 Jun 2023]
  -

- Author Correction
- Published: 01 June 2023

Author Correction: Single-neuronal predictions of others' beliefs in humans

- Mohsen Jamali,
- Benjamin L. Grannan,
- Evelina Fedorenko,
- Rebecca Saxe,
- Raymundo Báez-Mendoza &
- …
- Ziv M. Williams

Show authors

Nature **volume 618**, page E25 (2023)

- 538 Accesses
- 1 Altmetric
- Metrics details

The Original Article was published on 27 January 2021

Correction to: *Nature* https://doi.org/10.1038/s41586-021-03184-0
Published online 27 January 2021

When we recently reconstructed our recording locations on a standardized 3D brain model, we observed that the recording sites were in near proximity to those originally estimated but were also more posterolateral and somewhat broader in distribution. These recordings collectively spanned the superior frontal gyrus as well as part of its medial middle frontal gyrus border. These observations do not alter the neuronal findings or results of the paper and do not affect the analyses or other figures. There are two interesting implications of these observations, though, with respect to prior functional imaging studies.

First, the locations of dorsomedial prefrontal activities associated with theory of mind on neuroimaging generally lay anterior to the reconstructed recordings[1,2,3,4,5,6,7,8]. As a result, it seems likely that the neuronal populations involved in theory of mind, as reported here, are drawn from near[5,6,7,8] or outside[1,2,3,4] these selective regions. Indeed, as originally reported in the paper, we found a mixture of neurons within the recorded population that encoded social as well as non-social information, consistent with the relative lack of regional selectivity on imaging studies. Second, whereas the regions characterized by neuroimaging may play a domain-specific role in theory of mind inference, it is likely that the neuronal populations in the present study may be more centrally involved in theory of mind performance. Indeed, as reported in the paper, behavioral response accuracy was closely correlated with neuronal decoding accuracy when examined across task conditions. Here, the correction provides a detailed account of these recording locations.

In this correction, we include a new Supplementary Figure that illustrates the individual recording locations on a standardized 3D brain model in MNI152 space. We also include a new Supplementary Table that links the recording coordinates (SPM12) to each participant and provide an updated Supplementary Methods section that describes this procedure.

Supplementary information is available in the online version of this article.

References

1. Saxe, R. & Powell, L. J. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* **17**, 692–699 (2006).

2. Schurz, M., Radua, J., Aichhorn, M., Richlan, F. & Perner, J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42**, 9–34 (2014).

3. Carrington, S. J. & Bailey, A. J. Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum. Brain Mapp.* **30**, 2313–2335 (2009).

4. Dodell-Feder, D., Koster-Hale, J., Bedny, M. & Saxe, R. fMRI item analysis in a theory of mind task. *Neuroimage* **55**, 705–712 (2011).

5. Dohnel, K. et al. Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *Neuroimage* **60**, 1652–1661 (2012).

6. Sommer, M. et al. Neural correlates of true and false belief reasoning. *Neuroimage* **35**, 1378–1384 (2007).

7. Grezes, J., Frith, C. D. & Passingham, R. E. Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage* **21**, 744–750 (2004).

8. Hartwright, C. E., Apperly, I. A. & Hansen, P. C. Representation, control, or reasoning? Distinct functions for theory of mind within the medial prefrontal cortex. *J. Cogn. Neurosci.* **26**, 683–698 (2014).

Author information

Authors and Affiliations

1. Department of Neurosurgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Mohsen Jamali, Benjamin L. Grannan, Raymundo Báez-Mendoza & Ziv M. Williams

2. Brain and Cognitive Sciences, Massachusetts Institute of Technology, Boston, MA, USA

Evelina Fedorenko & Rebecca Saxe

3. Harvard-MIT Division of Health Sciences and Technology, Boston, MA, USA

Ziv M. Williams

4. Program in Neuroscience, Harvard Medical School, Boston, MA, USA

Ziv M. Williams

Corresponding author

Correspondence to Ziv M. Williams.

Supplementary information

Supplementary Figure

Supplementary Table

Supplementary Methods

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Jamali, M., Grannan, B.L., Fedorenko, E. *et al.* Author Correction: Single-neuronal predictions of others' beliefs in humans. *Nature* **618**, E25 (2023). https://doi.org/10.1038/s41586-023-06263-6

- Published: 01 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06263-6

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06263-6

- Publisher Correction
- Published: 02 June 2023

Publisher Correction: Crop switching can enhance environmental sustainability and farmer incomes in China

- Wei Xie ,
- Anfeng Zhu,
- Tariq Ali,
- Zhengtao Zhang,
- Xiaoguang Chen ,
- Feng Wu ,
- Jikun Huang &
- …
- Kyle Frankel Davis

Show authors

Nature **volume 618**, page E26 (2023)

- 468 Accesses
- 1 Altmetric
- Metrics details

The Original Article was published on 16 March 2023

Correction to: *Nature* https://doi.org/10.1038/s41586-023-05799-x
Published online 16 March 2023

In the version of this article initially published, a production error led to mistaken units being shown in the *x*-axis labels of Figure 2a,b, which are now amended in the HTML and PDF versions of the article.

Author information

Author notes

1. These authors contributed equally: Wei Xie, Anfeng Zhu, Tariq Ali

Authors and Affiliations

1. China Center for Agricultural Policy, School of Advanced Agricultural Sciences, Peking University, Beijing, China

   Wei Xie, Anfeng Zhu & Jikun Huang

2. School of Economics and Management, Jiangxi Agricultural University, Nanchang, China

   Tariq Ali

3. School of National Safety and Emergency Management, Beijing Normal University, Beijing, China

   Zhengtao Zhang

4. Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu, China

   Xiaoguang Chen

5. Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Feng Wu

6. Department of Geography and Spatial Sciences, University of Delaware, Newark, DE, USA

   Kyle Frankel Davis

7. Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA

   Kyle Frankel Davis

Corresponding authors

Correspondence to Wei Xie, Xiaoguang Chen, Feng Wu or Kyle Frankel Davis.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Xie, W., Zhu, A., Ali, T. *et al.* Publisher Correction: Crop switching can enhance environmental sustainability and farmer incomes in China. *Nature* **618**, E26 (2023). https://doi.org/10.1038/s41586-023-06245-8

- Published: 02 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06245-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06245-8

- Publisher Correction
- Published: 01 June 2023

Publisher Correction: Targeting PD-L2–RGMb overcomes microbiome-related immunotherapy resistance

- Joon Seok Park,
- Francesca S. Gazzaniga,
- Meng Wu,
- Amalia K. Luthens,
- Jacob Gillis,
- Wen Zheng,
- Martin W. LaFleur,
- Sarah B. Johnson,
- Golnaz Morad,
- Elizabeth M. Park,
- Yifan Zhou,
- Stephanie S. Watowich,
- Jennifer A. Wargo,
- Gordon J. Freeman ,
- Dennis L. Kasper &
- …
- Arlene H. Sharpe

Show authors

Nature **volume 618**, page E27 (2023)

- 567 Accesses

- 5 Altmetric

- Metrics details

The Original Article was published on 03 May 2023

Correction to: *Nature* https://doi.org/10.1038/s41586-023-06026-3
Published online 3 May 2023

In the version of the article originally published, in Figure 4a, the left- and right-hand labels now reading “Anti-PD-L2” originally read “Anti-PD-L1”. The labels have been corrected in the HTML and PDF versions of the article.

Author information

Author notes

1. These authors contributed equally: Joon Seok Park, Francesca S. Gazzaniga

2. These authors jointly supervised this work: Gordon J. Freeman, Dennis L. Kasper, Arlene H. Sharpe

Authors and Affiliations

1. Department of Immunology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

   Joon Seok Park, Francesca S. Gazzaniga, Meng Wu, Amalia K. Luthens, Jacob Gillis, Wen Zheng, Martin W. LaFleur, Dennis L. Kasper & Arlene H. Sharpe

2. Molecular Pathology Unit, Massachusetts General Hospital, Charlestown, MA, USA

Francesca S. Gazzaniga

3. Department of Pathology, Harvard Medical School, Boston, MA, USA

Francesca S. Gazzaniga & Arlene H. Sharpe

4. Program for Innovative Microbiome and Translational Research, Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Sarah B. Johnson, Golnaz Morad, Elizabeth M. Park, Yifan Zhou, Stephanie S. Watowich & Jennifer A. Wargo

5. Department of Immunology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Elizabeth M. Park, Yifan Zhou & Stephanie S. Watowich

6. Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Elizabeth M. Park & Jennifer A. Wargo

7. Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Jennifer A. Wargo

8. Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA

Gordon J. Freeman

Corresponding authors

Correspondence to Gordon J. Freeman, Dennis L. Kasper or Arlene H. Sharpe.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Park, J.S., Gazzaniga, F.S., Wu, M. *et al.* Publisher Correction: Targeting PD-L2–RGMb overcomes microbiome-related immunotherapy resistance. *Nature* **618**, E27 (2023). https://doi.org/10.1038/s41586-023-06237-8

- Published: 01 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06237-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06237-8

- Publisher Correction
- Published: 06 June 2023

Publisher Correction: Cortico-cortical feedback engages active dendrites in visual cortex

- Mehmet Fişek ,
- Dustin Herrmann,
- Alexander Egea-Weiss,
- Matilda Cloves,
- Lisa Bauer,
- Tai-Ying Lee,
- Lloyd E. Russell &
- …
- Michael Häusser

Show authors

Nature **volume 618**, page E28 (2023)

- 369 Accesses

- 1 Altmetric

- Metrics details

The Original Article was published on 03 May 2023

Correction to: *Nature* https://doi.org/10.1038/s41586-023-06007-6
Published online 3 May 2023

This paper was originally published under a standard Springer Nature license (© The Author(s), under exclusive licence to Springer Nature Limited). It is now available as an open-access paper under a Creative Commons Attribution 4.0 International license, © The Author(s). The error has been corrected in the HTML and PDF versions of the article.

Author information

Author notes

1. These authors contributed equally: Mehmet Fişek, Dustin Herrmann

Authors and Affiliations

1. Wolfson Institute for Biomedical Research and Department of Neuroscience, Physiology and Pharmacology, University College London, London, UK

   Mehmet Fişek, Dustin Herrmann, Alexander Egea-Weiss, Matilda Cloves, Lisa Bauer, Tai-Ying Lee, Lloyd E. Russell & Michael Häusser

Corresponding authors

Correspondence to Mehmet Fişek or Michael Häusser.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Fişek, M., Herrmann, D., Egea-Weiss, A. *et al.* Publisher Correction: Cortico-cortical feedback engages active dendrites in visual cortex. *Nature* **618**, E28 (2023). https://doi.org/10.1038/s41586-023-06241-y

- Published: 06 June 2023
- Issue Date: 22 June 2023
- DOI: https://doi.org/10.1038/s41586-023-06241-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06241-y